DOCUMENT RESUME

ED 431 209 EA 029 500

AUTHOR	Lashway, Larry
TITLE	Measuring Leadership: A Guide to Assessment for Development of School Executives.
INSTITUTION	ERIC Clearinghouse on Educational Management, Eugene, OR.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
ISBN	ISBN-0-86552-140-9
PUB DATE	1999-00-00
NOTE	128p.; Foreword by Kenneth Leithwood.
CONTRACT	RR93002006
AVAILABLE FROM	ERIC Clearinghouse on Educational Management, 5207 University of Oregon, Eugene, OR 97403-5207; Web site: http://eric.uoregon.edu; Tel: 541-346-5044; Fax: 541-356-2334 ($9.75).
PUB TYPE	Guides - Non-Classroom (055) -- ERIC Publications (071)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Administrator Effectiveness; *Administrator Evaluation; Educational Administration; Elementary Secondary Education; *Evaluation Methods; *Instructional Leadership; Management Development; Principals; *School Administration

ABSTRACT

This is a "how to" book for school-leader selection, appraisal, and development. It focuses on ways in which to measure components of leadership so that leadership evaluation can move beyond impressionistic opinions. The text is divided into five chapters. The rationale for leadership assessment, along with the dangers to be avoided in such assessment, are explored in the first chapter. Chapter 2 provides an overview of various perspectives on leadership. A selective review of the non-school-leadership literature is followed by a brief treatment of the school-leadership literature. Examples of instruments used to assess leadership, viewed from each of the different perspectives, are referred to throughout the chapter. Chapters 3 and 4 examine the interdependence of the needs of every school district, the forms of leadership best suited to meet those needs, and strategies for choosing a leadership-assessment instrument. These two chapters address how to select an instrument for a school district and then how to interpret and use the data generated by such an instrument. Technical issues are also examined. In the last chapter, nearly 20 instruments for assessing leadership are described, including their purposes, basic features, type of feedback garnered, and appropriate followup. (RJM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 431 209

MEASURING LEADERSHIP

MEASURING LEADERSHIP

MEASURING LEADERSHIP

LARRY LASHWAY

Foreword by Kenneth Leithwood

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ERIC

ERIC Clearinghouse on Educational Management

BEST COPY AVAILABLE

EA 028 750

MEASURING LEADERSHIP

MEASURING LEADERSHIP
MEASURING LEADERSHIP

A Guide to Assessment for Development of School Executives

LARRY LASHWAY

Foreword by
Kenneth Leithwood

ERIC

ERIC Clearinghouse on Educational Management

University of Oregon
1999

Library of Congress Cataloging-in-Publication Data

Lashway, Larry.

Measuring leadership : a guide to assessment for development of school executives / Larry Lashway ; foreword by Kenneth Leithwood.

p. cm.

Includes bibliographical references (p.).

ISBN 0-86552-140-9

1. Educational leadership—United States. 2. School management and organization—United States. I. Title.

LB2806.L28 1999

371.2'00973—dc21

99-19446

CIP

Design: LeeAnn August

Type: 11/12.5 Times

Printer: Cushing-Malloy, Inc., Ann Arbor, Michigan

E1683

Printed in the United States of America, 1999

ERIC Clearinghouse on Educational Management

5207 University of Oregon

Eugene, OR 97403-5207

Telephone: (541) 346-5044 Fax: (541) 346-2334

World Wide Web: http://eric.uoregon.edu

ERIC/CEM Accession Number: EA 029 500

This publication was prepared in part with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract no. OERI-RR 93002006. The opinions expressed in this report do not necessarily reflect the positions or policies of the Department of Education. No federal funds were used in the printing of this publication.

The University of Oregon is an equal opportunity, affirmative action institution committed to cultural diversity.

MISSION OF ERIC AND THE CLEARINGHOUSE

The Educational Resources Information Center (ERIC) is a national information system operated by the U.S. Department of Education. ERIC serves the educational community by disseminating research results and other resource information that can be used in developing more effective educational programs.

The ERIC Clearinghouse on Educational Management, one of several such units in the system, was established at the University of Oregon in 1966. The Clearinghouse and its companion units process research reports and journal articles for announcement in ERIC's index and abstract bulletins.

Research reports are announced in *Resources in Education* (*RIE*), available in many libraries and by subscription from the United States Government Printing Office, Washington, D.C. 20402-9371.

Most of the documents listed in *RIE* can be purchased through the ERIC Document Reproduction Service, operated by Cincinnati Bell Information Systems.

Journal articles are announced in *Current Index to Journals in Education*. *CIJE* is also available in many libraries and can be ordered from Oryx Press, 4041 North Central Avenue at Indian School, Suite 700, Phoenix, Arizona 85012. Semiannual cumulations can be ordered separately.

Besides processing documents and journal articles, the Clearinghouse prepares bibliographies, literature reviews, monographs, and other interpretive research studies on topics in its educational area.

CLEARINGHOUSE NATIONAL ADVISORY BOARD

Terry A. Astuto, Vice-President, Division A, American Educational Research Association
George Babigian, Executive Director, American Education Finance Association
Anne L. Bryant, Executive Director, National School Boards Association
Patrick Forsyth, Executive Director, University Council for Educational Administration
Paul Houston, Executive Director, American Association of School Administrators
John T. MacDonald, Director, State Leadership Center, Council of Chief State School Officers
Philip K. Piele, (Chair), Professor and Director, ERIC Clearinghouse on Educational Management, University of Oregon
Samuel G. Sava, Executive Director, National Association of Elementary School Principals
Gerald Tirozzi, Executive Director, National Association of Secondary School Principals

ADMINISTRATIVE STAFF

Philip K. Piele, Professor and Director
Stuart C. Smith, Associate Director

Contents

Foreword vii

Introduction 1

Chapter 1: **The Measure of a Leader** 5
- The Search for Measurement 7
- The Case for Systematic Measurement 11
- The Dangers of Measurement 14
- How Is Leadership Tested? 16

Chapter 2: **What Do Leadership Tests Measure?** 23
- Looking at Leadership 23
- Leadership Qualities 37
- The Need to Choose Wisely 42

Chapter 3: **Selecting the Best Instrument** 45
- Step 1: Know Your Purpose 45
- Step 2: Develop a Local Definition of Leadership 47
- Step 3: Locate Possible Instruments 50
- Step 4: Determine the Suitability of Various Tests for Your Needs 51
- Step 5: Establish the Credibility of the Instrument 52
- Step 6: Consider Practical Issues 58
- Step 7: Consider the Ethical and Legal Issues 58
- Step 8: Determine How Much Support the Instrument Provides for Followup and Professional Development 59

Chapter 4: **Fulfilling the Promise of Assessment** 61
- Selecting the Facilitator 62
- Deciding on the Setting 63
- Delivering the Feedback 65
- Forming a Development Plan 75
- Searching for Ground Truth 77

Chpater 5: **Instrument Profiles** 81
Benchmarks 81
COMPASS: The Managerial Practices Survey 83
The Comprehensive Leader 85
Educational Administrative Style Diagnosis Test (EASDT) 87
Educational Administrator Effectiveness Profile 89
Leader Behavior Analysis II (LBAII) 90
Leadership Competency Inventory (LCI) 91
Leadership/Impact™ 93
Leadership Orientations 95
Leadership Practices Inventory (LPI) 96
Leadership Skills 98
Leadership Sphere Inventory (LSI) 99
Leadership Workstyles 101
Life Styles Inventory (LSI) 102
Myers-Briggs Type Indicator (MBTI) 104
NASSP Assessment Centers 105
The Principal Perceiver 107
Professional Development Inventory 108
Profile of Aptitude for Leadership (PAL) 111
The Survey of Leadership Practices 112
The Survey of Management Practices 114

Bibliography 117

Foreword

Site-based management, high-stakes student testing, professional teaching standards, school choice—these are a few examples of the many educational reform initiatives that have aimed to increase the accountability of schools over the past decade and a half. For those in educational leadership roles, one consequence of this reform "zeitgeist" has been the expectation that their decisions will be data driven as never before. It seems fitting, in such a context, that decisions about leaders themselves also be more data driven. This concise and highly readable book by Larry Lashway offers practical advice and resources toward that end.

Intended primarily for those with districtwide responsibilities, this is a "how to" book for school-leader selection, appraisal, and development. It is written extremely well with this audience in mind. And while the author touches, of necessity, on only a sample of the total set of leadership instruments that are available, he provides quite useful guidance for district administrators no matter which leadership instruments they select for their own purposes.

The rationale for leadership assessment, along with the dangers to be avoided in such assessment, are explored in the first chapter. Examples sprinkled throughout the chapter give practical meaning to these issues, which all too often have been presented only in abstract terms.

An overview of various perspectives on leadership is found in the second chapter. A selective review of the non-school-leadership literature is followed by a brief treatment of the school-leadership literature (nothing is said about instructional leadership, however). Examples of instruments used to assess leadership viewed from each of the different perspectives are referred to throughout the chapter. One of the main purposes for this and the next two chapters is the seemingly obvious but often overlooked interdependence of the needs

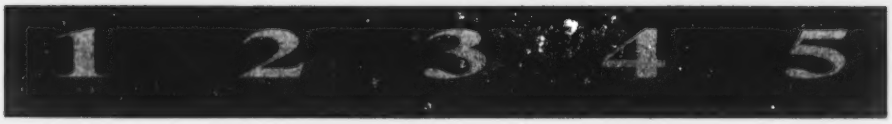
of one's district, the forms of leadership best suited to meet those needs, and one's choice of leadership-assessment instrument.

Chapters 3 and 4 address two related issues: how to select or customize an instrument for your own district (chapter 3); and how to interpret and use the data generated by such an instrument to improve leadership in your own district (chapter 4). Most of the technical issues of which one needs to be aware in selecting and using primarily quantitative, survey-type instruments to assess leadership are examined in these chapters. Lashway does this job exceptionally well, offering highly accessible accounts, with illustrations, of such matters as instrument validity and reliability.

A selection of nearly twenty instruments for assessing leadership, most available from commercial sources, are described in the fifth and final chapter of this brief book. For each instrument, this description consists of its purposes, basic features, nature of the feedback provided by the instrument, and the followup recommended in response to such feedback. Information is also provided about the theory (if any) on which the instrument is based, how to administer the instrument, statistical validation data about the instrument (if available), uses, costs, and address for more information about the instrument.

This book fills an important niche in educational measurement. Lashway speaks directly to the practical information needs of busy senior administrators. He does this by stripping away the often obtuse language used to present information that is, nonetheless, extremely important for school executives to consider in their decision-making. Furthermore, his book aims to bring greater rigor to an area of their decision-making that has, arguably, the greatest long-term impact on their districts.

Kenneth Leithwood, Ph.D.
Director and Professor
Centre for Leadership Development
Ontario Institute for Studies in Education
University of Toronto



Introduction

Americans have long had a love-hate relationship with testing. As a nation that values a practical, no-nonsense approach to daily life, we are intrigued by the notion that thirty to sixty minutes with paper and pencil can cut through our foggy perceptions to provide a reliable, valid measure of the qualities we cherish, from intelligence to vocational aptitude. Popular magazines learned long ago that a surefire way to capture readers' attention is to offer a simple ten-minute quiz that claims to reveal readers' personality, emotional intelligence, or marital compatibility.

At the same time, we've learned to be suspicious of instruments that promise so much yet often deliver so little. Virtually all of us have had the experience of taking a test at school that somehow failed to capture the knowledge in our heads. If that can happen on a simple academic test, what can we expect from instruments that purport to measure intelligence, personality, or leadership?

Thus, when the ERIC Clearinghouse on Educational Management asked me to prepare a book with the title *Measuring Leadership*, I accepted the assignment with some ambivalence. While I knew enough about testing to respect the ability of a well-designed instrument to capture subtle aspects of human behavior, I had also spent enough time in schools to appreciate the complexity of the K-12 environment. I wondered whether any test could pinpoint the qualities needed to lead such unruly institutions.

But as I read the testing literature I found evidence that formal assessment could provide insights into dimensions of leadership that might otherwise go unnoticed. As I examined numerous instruments (and took advantage of the opportunity to assess my own leadership qualities), I was impressed that so many of them are thoughtful, well-designed, and relevant to school leadership (even when they have not been designed for that environment). In the process, I learned some

things about my own approach to leadership, and was stimulated to think about how it could be improved. My exploration into leadership assessment left me with little doubt about its potential value.

Yet my research also underscored the hazards of taking test results out of context. It was clear that these tests were not tapping into some universal leadership "essence." Each instrument uses different definitions and strategies to assess leadership qualities, and each offers a different lens through which leadership can be viewed. Simply learning the score is of little value unless the underlying model is understood. Productive leadership assessment requires careful attention both before and after the test has been administered.

Thus, this book focuses less on detailed descriptions of instruments than on a *process* that begins with reflection on the district's leadership needs and ends when participants begin to act on the implications of the results. My hope is that both district officials and principals will find practical guidance in this discussion.

As with most ERIC publications, this volume is a work of synthesis, designed to report on the existing literature rather than to create new theories. However, any synthesis inevitably requires personal judgments and interpretations. In the attempt to formulate a coherent picture from diverse sources, I may have seen implications or made connections that the original authors did not intend. Thus, readers should not regard the material in these pages as the final word on the subject, but only as an initial recommendation that should be freely challenged and adapted through their own insights and experiences.

I want to express my appreciation to the Clearinghouse for offering me the opportunity to write this volume. I owe a special debt to Stuart Smith, whose patience I severely tested but whose support was unflagging. Additional thanks go to Mike Krigelski, superintendent of schools in Centreville, Michigan, whose commonsense comments put a human face on the testing process.

As always, a work of this type relies on access to good libraries, and I've been blessed with an abundance of rich resources, including The Evergreen State College, Pacific Lutheran University, the Washington State Library, and Olympia Timberland Library. Their assistance made a difficult task much easier.

FORETASTE

Since ancient times, people have searched for ways of identifying leaders who would be strong enough, brave enough, or smart enough to guide them through perilous times. In today's schools, that need is stronger than ever.

While many schools continue to use informal methods of assessing leadership, some are moving toward systematic measurement of leadership qualities. Through the thoughtful use of well-designed instruments or activities, schools can add depth, breadth, and objectivity to their assessment of administrators.

Casual or careless use of measurement instruments, however, can give misleading results. Schools must choose carefully from the wide variety of tools available, ranging from paper-and-pencil tests to performance assessments.

This Chapter's Contents:

The Search for Measurement
The Case for Systematic Measurement
The Dangers of Measurement
How Is Leadership Tested?

The Measure of a Leader

When the great King Uther died, the kingdom fell into great peril, for he had no heir and many wished to rule the land. The Bishop called the lords together to pray and await some divine sign saying who should be king. And when they had done so, they found in the churchyard a great stone in which was embedded an anvil, and in the middle of the stone and anvil was a sword, on which was written: Whoso Pulleth Out This Sword From This Stone and Anvil, Is the True and Rightful King. Many great lords took their turn at trying to wrestle the sword out, but none could move it the least bit. Then came the boy Arthur and easily pulled out the sword, which was called Excalibur. Seeing this, the great lords knelt and recognized him as king.

—English legend

Modern organizations can only envy the denizens of Camelot, who could count on divine intervention to anoint their leaders. Arthur proved his worth with a task that was simple, heroic, and unambiguous.

Today the challenge of identifying leadership qualities seems complex and confusing. Consider the following scenarios:

As the last candidate left the interview room, Marcia Trulo leaned back, thought for a moment, and announced a 10-minute break. After a full day of interviewing applicants for the Meadowdale position, her Executive Council needed time to stretch and refresh—and she needed time to think.

On paper the five candidates all looked good—not much to separate them. One of them had a rather shaky interview, but the other four all came across well. They had differing patterns of strengths and weaknesses, but overall none was head and shoulders above the rest. On that analysis, perhaps any of the four would be good choices. But she could recall some other new hires whose performance failed to match the impression they created during the interviews.

Given all that was happening at Meadowdale—changing boundaries, high faculty turnover, and declining test scores—they couldn't afford an unlucky guess. As she thought back on the interviews, she tried to identify some sign—however small the clue—that this was the person with the necessary leadership qualities.

Mark Lee sat at his desk staring out the window. The students had left two days ago, the last teacher had turned in keys an hour ago, and his end-of-year reports weren't due for a week, so it was a rare moment for reflection. He couldn't believe how fast his first year as principal had gone, and at this point he wasn't even sure *how* it had gone.

His university classes had taught him some good techniques for getting things done, but they hadn't told him much about assessing his effectiveness. He had felt off-balance all year; projects he had been really nervous about went off without a hitch, while seemingly benign issues had suddenly blossomed into contentious debates.

Overall, things had moved forward a little, but Lee hadn't yet convinced himself that he had the leadership qualities to take this school where it needed to go. He knew there were some things he needed to improve on, like organization, but he just didn't have an overall sense of direction.

Both Trulo and Lee are looking for a reliable way to measure leadership. Trulo wants to determine which candidate has enough of the right qualities; Lee wants to know how to improve his own qualifications and skills.

Why should this be difficult? Anyone who has worked with other people knows that leadership exists. There are people of whom we can say, "He's a leader," and everyone nods in agreement. Yet when

we try to pinpoint exactly what makes us feel that way, we bog down. Scholars haven't done much better, generating hundreds of definitions but little consensus.

Leadership appears to be one of those concepts (much like "love" or "intelligence") that plays a vital role in our vocabulary but that stubbornly resists distillation into a pure essence. Leadership flows from many sources, sometimes springing from the joy of accomplishment, other times from a modest desire to serve others. It takes many forms, sometimes visible and heroic, other times quiet and unassuming. It has different effects in different environments; a strategy that succeeds brilliantly in one organization may completely fail in another.

Given all that, is there hope for Trulo and Lee? Can we move beyond raw intuition and gut-level hunches to assess leadership systematically and reliably? The answer is a qualified yes. While there is no Excalibur test—no heroic task to reveal divine favor—there are many tools that can help those who wish to choose leaders or assess their own leadership qualities. This book is designed to provide an introduction to the measurement instruments that can help school leaders.

The Search for Measurement

Edward Thorndike, the great educational psychologist, once summed up his worldview by saying, "Whatever exists at all exists in some amount." In principle, everything—no matter how abstract or complex—could (and should) be measured. The role of scientific psychology was to find objective, reliable yardsticks for human behavior.

Thorndike was a towering figure in his time, and his optimistic views permeated the educational establishment, spurring several generations of educators to validate their work with numbers. Tests were designed to assess intelligence, academic achievement, musical ability, handwriting, and dozens of other qualities. The corollary to Thorndike's credo soon became "If it exists, someone will try to measure it."

Today educators are more skeptical of "scientific objectivity," but the desire to capture elusive qualities in numerical form remains strong. Nowhere is the need more strongly felt than in the leadership domain.

High Stakes

Schools have always needed strong leadership, but the stakes have never been higher than they are today. Consider:

- Demographic and cultural shifts are creating increasing numbers of the students that schools have traditionally been least successful in helping.
- Social changes are sending students to school with needs that go far beyond the academic.
- Public expectations have risen, demanding that schools not only help students reach higher standards but that they be sensitive to individual differences in style.
- A growing number of critics have suggested that the public education system as a whole should be scrapped; radical reforms such as voucher plans now receive more serious support than ever before.
- Even defenders of public education now agree that serious restructuring is in order.

Several decades ago, school boards could expect satisfaction if they hired leaders who were congenial, hard-working, and organized enough to get their paperwork done on time. Whatever shortcomings these leaders had could be easily smoothed over, and no one doubted that school would continue to run in much the same way it always had.

In today's volatile environment, schools pay a much higher price for a poor choice. If a new principal does not have the necessary skills, progress grinds to a halt while the entire school tries to cope with the resulting confusion, disappointment, and conflict. Often the failing leader becomes *the* dominant issue in the minds of teachers and parents. Equally damaging is the new leader who is just competent enough to avoid overt trouble, but who does not take the school in any particular direction. Years may pass before his inaction or indecisiveness suddenly explodes into a crisis.

The stakes are equally high for professional development. No preparation program can prepare leaders for everything they will encounter in their careers. Today's leaders must be learners, swiftly making adjustments to a rapidly changing environment. Without accurate feedback about their performance, their self-improvement efforts will falter.

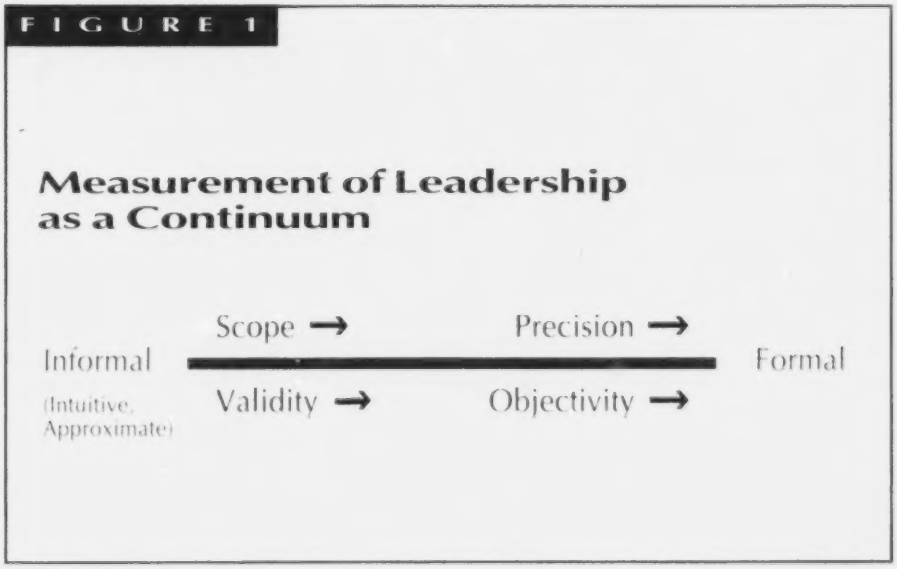
The Measurement Continuum

When leaders and their supervisors become dissatisfied with informal or intuitive appraisals, they usually look for some kind of test that will identify leadership qualities in tangible, precise ways. What they envision is what we have come to associate with testing: an instrument that will reduce complex behaviors to a set of numbers that can be objectively ranked, compared, and analyzed. In this common view, the choice is between measuring or not measuring.

However, the reality is more complicated. Measurement actually exists on a continuum (see figure 1). At one end, measurement is intuitive, impressionistic, and inexact. We may merely decide that one person has "more" of the desired quality than another, but this is still a measurement. (The old phrase "taking his measure" captures this idea, suggesting the almost instinctive way that people size up others, trying to locate their place on some scale of skill, courage, or other quality.)

At the other end of the continuum, leadership is assessed through systematic data in numerical form. The goal is to render a judgment that is less influenced by surface appearances and human biases.

FIGURE 1



As we move toward the right side of the continuum we gain four advantages.

1. *Objectivity.* Objectivity means that the same criteria are applied in the same way to everyone. For example, on an objective

assessment, each candidate is asked the same questions, and the same criteria are used to rate each response. When objectivity is present, test users are assured they aren't comparing apples to oranges. In contrast, the typical informal interview is idiosyncratic, a dialogue that wanders off in unpredictable directions. One applicant may get favorable marks for her response to a question that another applicant isn't even asked.

2. *Validity.* Validity means that the criteria being used have a relationship to the quality being assessed. A candidate may give a seemingly apt response to a question in an interview, but does that response have any relationship to later success on the job? Formal tests frequently address that question in systematic ways. For example, given the scores on a leadership test, we can locate the subjects several years later and see whether their scores on the test relate to their success on the job. Do high scorers tend to be judged as outstanding performers in the field? Do low scorers encounter difficulties on the job? Such studies are difficult and expensive to carry out (and not every test goes to this length), but most formal measures have been validated in some way.

3. *Precision.* Precision allows us to make fine distinctions rather than crude estimates. For example, a principal doing an informal self-assessment may recognize that she is good at resolving conflicts, but a formal test may tell her that she scores at the ninety-ninth percentile in conflict resolution. The difference? The quantitative score may lead her to see this skill as a unique strength that should be cultivated and expanded. On most formal tests, the precision is particularly valuable because it allows us to compare leaders with a much broader reference group than we could find at the local level. That is, we can know how a leader rates among a representative national sample, not just the handful of leaders we know personally.

4. *Scope.* Formal tests generally cover a broader spectrum of qualities and skills more efficiently than we could do informally. For example, the *Educational Administrator Effectiveness Profile* (Human Synergistics) offers 120 items that tap 11 key leadership skills in less than an hour. Because each skill area is probed with multiple questions, there is little chance that any single response can dominate the outcome (as may happen with unstructured interviews).

Viewing measurement on a continuous scale helps us recognize a crucial point: *all* assessment involves human judgment. At one end of the scale, the judgment is informal, intuitive, and approximate; at the other end it is systematic, objective, and precise—but only relatively so. No matter how sophisticated the calculations or how extensive the

validation studies, every test embodies human judgments about leadership and how it can be measured. That is, at some point, the creators of the test have faced—and answered—the crucial questions, "What is leadership?" and "How can it best be measured?" Every test answers these questions in different ways, and every test will mean something different than any other test.

The Case for Systematic Measurement

In many school districts, leadership is measured on the left end of the continuum. For example, the hiring process begins with a screening of résumés in which screeners match the materials against some list of desired attributes (implicit or explicit). Once the candidates have been whittled down to three to five, they are brought in for interviews, which are frequently loose and improvisational, often opening with, "Tell us a little about yourself" and meandering along an unpredictable path. The result? Leaders are chosen on impressions and idiosyncratic perceptions of "fit" rather than merit (Mark Anderson 1991).

Some districts have attempted to move this process further to the right on the measurement continuum. Before reviewing résumés, they explicitly list desired criteria and then check off the ones they find on each résumé. In setting up interviews, they identify specific issues that should be addressed and make sure that each candidate is asked those questions. Sometimes the interviewers rate the answers numerically. Throughout the process, some schools attempt to reduce the effect of subjectivity by involving multiple parties in the selection process (Anderson).

The same picture is apparent in the professional development of school leaders. In many cases, leaders are left to ponder their strengths and weaknesses on their own, based on whatever day-to-day feedback they've picked up from teachers, parents, and other administrators. They may be evaluated for accountability purposes, but administrators get little systematic feedback that could be used for professional development (Anderson).

Here again, however, some districts have attempted to move toward more systematic measurement. A number of districts have implemented carefully designed assessments that provide "360 degree feedback" in which leaders are assessed by superiors, subordinates, peers, and others in the school environment (Richard Manatt).

In summary, much of leadership assessment has been unstructured and informal, but some districts are seeking to make their

practices more systematic and precise. Are these districts on the right track? Will formal measurement of leadership yield richer, more accurate information? The answer is "yes."

Informal or intuitive judgments can be easily influenced by extraneous factors—physical appearance, self-confidence, and apparent fit with community values (Anderson). What one sees in an hour-long interview is exactly one hour of behavior, mostly verbal, carefully crafted to create a certain impression. Certain qualities and skills may fail to surface altogether (for example, a candidate's ability to build and maintain relationships with teachers).

Mike Krigelski, superintendent of schools in Centreville, Michigan, uses the *Principal Perceiver* (Gallup Organization) as part of the selection process, because it picks up strengths in candidates who may not interview well. He recalls one applicant who "didn't have the knockout look," but who turned out to be caring, dedicated, and appreciated by students—qualities the *Perceiver* recognized. Noting that interviews are often a matter of impression management, Krigelski wryly observes, "I trust my gut, but sometimes it turns out to be just gas."

Self-assessment (the most common form of assessment for professional development) is also notoriously susceptible to bias and blind spots. Leaders' self-assessments are frequently more positive than that of superiors and subordinates (Bernard Bass).

Kenneth and Miriam Clark (1996) argue that information that is collected systematically and combined objectively provides better predictors of performance than observer judgments—even when the observer is highly qualified and knows the candidate well.

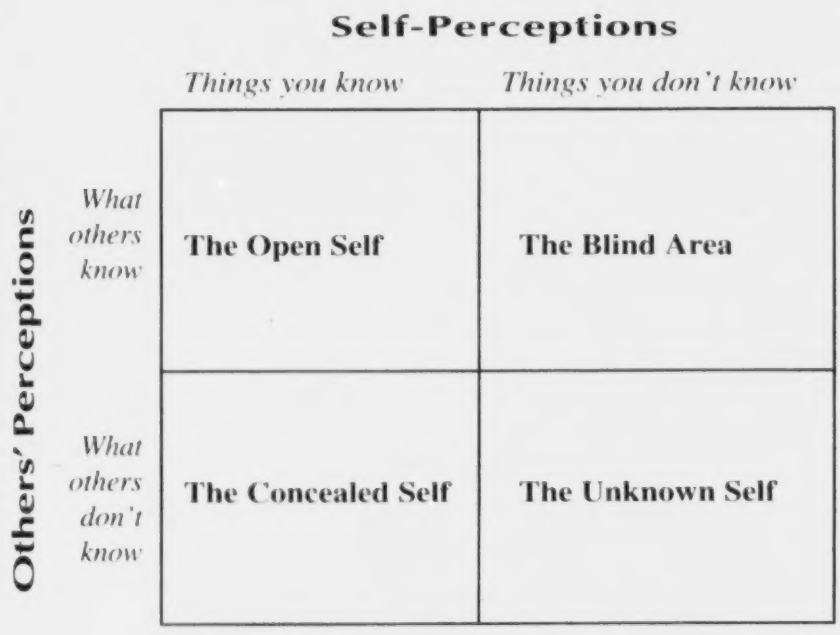
The benefits of systematic measurement can be seen more clearly by using the Johari window, a tool frequently used for analyzing interpersonal communication (Hanson).

The Johari window, displayed in figure 2, simply shows the different ways that people's skills, values, and personal qualities can be known or not known to themselves and others. Each quadrant of the window represents a different degree of knowledge of an individual by self and others.

1. *Qualities known to an individual and to others are part of the "open self."* For example, a leader sees herself as being "supportive," and those who know her agree. In this case, a leadership test seems superfluous, since everyone is already aware of the leader's characteristics. But a credible formal test can confirm the congruence

FIGURE 2

The Johari Window



between self-perception and others' perception, elevating the quality from "something I can do" to "something that is a real strength."

2. *Qualities known to others but not to the individual are part of the "blind area."* Here is where formal feedback from others carries real value. For example, a principal may be completely unaware that certain behaviors are creating an aura of "aloofness"; getting that feedback in a structured way can serve as a wakeup call, raising all kinds of productive questions about why this is the case and what can be done about it.

3. *Qualities known to the individual but not to others are part of the "concealed self."* A leader may present the appearance of an organized, efficient manager yet be aware that behind the scenes his use of time is often unproductive or undisciplined. A good test may be useful in two ways in such a case. First, a skillfully designed

instrument may draw out responses that the leader would not directly state to others; what has been concealed is now revealed. Second, seeing in black and white that one's facade has been successful may make a leader uncomfortable enough to do something about it.

4. *Qualities that are unknown to both the self and others are in "the unknown self."* At first glance, it seems as if tests would be of little help here; if people aren't aware of certain qualities, how will the test pick them up? Part of the answer is that tests often focus on concrete behaviors that *are* known, but then go on to establish connections that reveal larger themes. For example, a principal may routinely make a point of returning all phone messages before leaving each day, but not recognize this as an indicator of skill in interpersonal communications. The test will point that out.

In addition, viewing the results of a test may encourage leaders to reflect on their strengths and weaknesses, sometimes leading to discoveries that the test per se did not reveal. For example, a leader pondering high scores in "dominance" or "directiveness" may suddenly realize that those behaviors are not a consciously chosen strategy but an echo of his relationship with his father.

Users of the Johari window generally believe that performance and relationships improve when the open self expands and the other areas shrink. That is, when a leader's qualities are known by the leader as well as by others, the leader will function better (Hanson). Formal tests can play a major role in that process.

The Dangers of Measurement

A well-designed program for measuring leadership is a valuable resource for choosing and selecting leaders, but testing is not a panacea, nor is it something that should be done casually. Numerous obstacles lie in wait for unwary school districts.

1. *Poorly designed tests.* Most tests generate numbers, scales, and even charts, creating an authoritative aura that seems to be above reproach. Although well-designed tests do provide good information that cannot be achieved in other ways, not all tests are well designed. Appropriate use requires careful investigation of the test's makeup, development, and statistical validity. This is an uncomfortable issue for many users, who are not adept in statistics, but there are some relatively simple steps that will help even nonexperts sort through the issues. Chapter 3 discusses these issues in more detail.

2. *Inappropriate tests.* As noted earlier, every test embodies some theory about the nature of leadership, and no test is completely comprehensive. Users must carefully evaluate their needs. How is leadership defined in their district? What are the most important leadership qualities needed for success in the local environment? Which test best matches the district's needs? No matter how well designed the test, its validity suffers when the qualities it measures are irrelevant in the district.

3. *Poor use.* Testing is never just a numbers game. The scores generated by any test are incomplete; they must also be analyzed, interpreted and put into context. For example, a test may show that Candidate A is at the ninety-fifth percentile while Candidate B is at the ninetieth percentile. In practical terms, is this difference significant? Does the five-point difference promise a major difference in on-the job-performance? Does Candidate A's superiority on the test outweigh the achievements on Candidate B's more substantial résumé? Test scores are an invitation to reflection and human judgment, not a substitute for it.

The same reasoning applies when tests are used for professional development. Simply reporting results to test-takers will be of little use unless there is a clear pathway for following up. Are leaders provided with accurate interpretations of their scores? Are they encouraged to reflect on the results in a systematic way? Are they provided resources for taking action on the issues that have been identified?

4. *Ethical risks.* In some respects, testing is a dehumanizing process. It takes the rich thought patterns of the leader and attempts to abstract a few preselected categories. Test-takers see their complex daily reality—the insights, intuitions, wisdom, and deeply felt values—reduced to a set of numbers. The experience can be unsettling, particularly in a high-stakes setting, where jobs and promotions might be involved. Robert Kaplan and Charles Paulus (1994) note that while most leaders can handle the stress of feedback, there is always the chance of "pain." Pain is not always harmful, but neither should people be subjected to it arbitrarily.

Consider the following scenarios:

- A candidate for a principalship, with an excellent record of achievement, is not hired because of a mediocre score on a leadership test that was hastily chosen by the school "because we want to be objective."

- An enthusiastic first-year principal is dismayed to find that teachers have rated him as "average" on a majority of items; the feedback results have simply been placed in his mailbox with no further explanation.
- The school board decides that the district's 360-degree-feedback results will be used to determine salaries and promotions; the instrument has been developed locally, with no field testing and validation.

In all these cases, careers may be damaged because districts have exercised little care in choosing and using test instruments. At a minimum, test users must have a clear, defensible purpose in administering a test and must use the results in an appropriate and productive way.

How Is Leadership Tested?

Leadership assessment can take many forms, each of which has advantages and disadvantages. This section discusses the range of tools available to schools.

Intuitive Judgments

Measurement experts have often criticized the use of subjective judgments (as in an unstructured interview), and they can point to evidence that assessors may be unduly influenced by peripheral factors that have little to do with long-range leadership effectiveness (Anderson).

Yet the persistent use of such judgments suggests that practitioners have found them to be a source of useful information. Even administrators who use formal assessments would be reluctant to hire someone sight unseen, based only on a set of test scores. So it may be worth asking what assessors can glean from an unstructured face-to-face meeting.

One possibility is demeanor. How do candidates present themselves? Are they good listeners? Do they speak with assurance and appear to be comfortable with themselves? Do their appearance and manner say "leader"? These may appear to be nebulous or even superficial qualities, but school leadership requires intensive interpersonal interactions; candidates who can present themselves persuasively in an interview may be able to do the same in other situations.

Philosophical compatibility is another useful piece of information that can come from interviews. Either through targeted questions or spontaneous discussion, assessors can often get glimpses of a candidate's worldview, allowing predictions about potential harmony or dissension. (Some critics point out the danger of "groupthink," in which members of an organization continually seek to clone themselves, screening out new ideas and stifling creativity. While that risk is real, few districts would find it sensible to hire someone who would try to turn the organization 180 degrees from its chosen path.)

In general, intuitive judgments can be useful in determining "fit." Without interviews, for example, we could anticipate many more cases of leaders getting into personality conflicts or philosophical disputes. Where informal assessments fall short is in assessing "merit"—the actual skill level of the candidate. Being able to talk about one's strengths, or even relating a story that illustrates those qualities, is a far cry from actually demonstrating the skills.

Performance Judgments

Intuitive judgments require a leap of faith that the behavior we see in the interview can be translated into effective behavior on the job. A more reliable judgment might come from observing how a person actually performs in a school setting. Will that pleasant smile evaporate when confronted with a demanding parent? Will the well-rehearsed recitation of educational vision turn to mush in the give and take of a faculty meeting? Observing on-the-job performance can answer many of those questions.

Performance assessments generally come from four sources.

1. *Superiors.* The most common form of on-the-job observation comes from superiors, who are generally assumed to be senior, more knowledgeable, and more highly skilled than the people they supervise. Superiors are especially able to provide feedback on the leader's ability to accomplish assigned tasks. In some settings (particularly smaller organizations), they may be able to provide information on how the leader interacts with peers and subordinates.

2. *Peers.* People occupying similar roles can often add insight into a leader's performance. From their own experience, they know what the job requires, and they can often get beneath the surface behavior to make thoughtful inferences about their colleague's thought

processes. Some studies have suggested that peer ratings may be more predictive than ratings by superiors (Bass).

3. *Subordinates.* Leaders, according to one popular definition, are people who have followers. Thus, it makes sense to tap followers' perceptions. Those on the receiving end of the leader's actions often have a keen awareness of what is and is not working, and they may be highly perceptive about the reasons.

4. *Self.* Many leaders are capable of rating their own performance honestly and frankly (if not objectively). In particular, leaders know better than anyone *why* they do what they do. They have a rich knowledge of their intentions, hopes, and values, and they can explain the thinking behind their strategies. However, they may not always recognize the way their actions are perceived by others; some research shows a wide gap between self-assessments and the assessment of others (Bass). Yet self-assessments can be useful as a starting point that will place the feedback of others into an understandable context.

While observations of job performance can provide excellent insight into a leader's skills, they also carry a number of disadvantages.

First, every human being sees the world through a particular "filter" that reflects his or her values, experiences, and aspirations. That filter not only affects our interpretation of another's behavior, it may even keep us from seeing things that would be obvious to someone with a different filter. For example, someone with a bent for brisk efficiency (a single-minded focus on the task at hand) will be likely to notice and value that trait in a subordinate, while his radar completely fails to pick up another subordinate's gift for warm, positive relationships.

Second, superiors, subordinates, and peers may have limited perspective. No one is likely to have a comprehensive picture of a leader's performance. Superintendents, for example, may be able to comment perceptively on how well principals communicate with the central office but know less about their communication with faculty. Similarly, teachers may be highly aware of the principal's behavior in a few key areas (such as supporting them on discipline issues), but be only vaguely aware of how he or she is handling other leadership tasks. (In such cases, ratings may depend more on the raters' state of mind than on careful analysis of the leader's behaviors. If teachers are happy, they may give high marks on most questions on the grounds that the principal must be doing *something* right.)

Third, assessments that have high stakes (promotion, hiring, and so forth) can be affected by micropolitical issues (Bass). A superintendent angling to remove a principal may rate her lower on many items to help make the case for a dismissal. Teachers may be dissatisfied with some parts of a principal's performance but rate him high because they know they can at least live with him.

These disadvantages are dangers, not fatal flaws. To some degree, they can be countered by providing multiple ratings from different perspectives. A growing trend is the use of "360-degree evaluation," which uses feedback from all groups—superiors, peers, subordinates, and sometime others (Manatt, Craig Chappelow).

Simulated Performance

Assessment of on-the-job performance is an especially helpful tool in professional development, since superiors, subordinates, and peers are readily available to give their opinion. However, it is less practical when selecting new leaders from outside the organization.

One alternative is measuring performance through simulated work activities. The best-known example is the "in-basket" exercise, in which candidates are given a list of typical action items (such as phone messages, memos, reports) that might appear on a leader's daily agenda, and asked to handle them just as they would in real life (Bass).

Such simulations can be done informally on the local level, or they can be expanded into a formal set of procedures carried out at an "assessment center." In the best known educational example, the National Association of Secondary School Principals runs centers around the country that assess leaders or potential leaders on a variety of skills.

Paper-and-Pencil Tests

What people most often associate with the term "test" is a written instrument containing anywhere from a dozen to a hundred statements or questions related to leadership behavior or beliefs. For example, test takers may be asked to agree or disagree with statements about their behavior such as:

- I frequently walk around the building just to stay in touch.

- I never make a decision without consulting those who will be affected.
- I frequently delegate tasks to others.
- I go out of my way to compliment people for good work.
- I attempt to return every phone message the same day.

Or they may be asked to report on their beliefs:

- Most of the time, employees can be trusted to do the right thing.
- In today's volatile environment, long-term planning makes little sense.
- Maintaining good communications is the leader's most important task.
- Good leaders are born, not made.
- It's better to introduce change a little at a time rather than all at once.

Tests of this kind yield at least two helpful bits of information. First, the responses can be grouped thematically to provide a profile of the leader's beliefs or behavior in a particular area. For example, *The Comprehensive Leader* (HRDQ) has forty items that can be grouped in four areas: "know yourself," "know others," "know your organization," and "know the world."

Second, the leader's score can be put in context by comparing it with scores of others who have taken the test. The score might be reported, for example, as being at the ninetieth percentile or in the top quartile of a nationwide sample of leaders who took the test. Even more helpfully, the scores might be compared to the scores attained by leaders with demonstrated success in the workplace; then a high performance on the test could be viewed as a predictor of future success.

In addition, paper-and-pencil tests tend to be time-efficient, seldom requiring more than an hour, and easily scored, either by sending them to the vendor or by using a self-scoring key. Users can quickly get structured, comprehensive profiles of a large number of leaders.

However, tests require considerable time and expertise to develop, something beyond the resources of most school districts. In

most cases, districts must choose a ready-made instrument and order it from a vendor.

In summary, school districts have a wide array of choices in assessing leadership, ranging from the informal to the highly structured. These choices are not mutually exclusive; using multiple approaches is possible and probably desirable (Clark and Clark).

Unfortunately, schools do not always make the best use of these resources. Kenneth Leithwood and colleagues (1994) cite "compelling evidence" that "information generated during appraisal and selection processes is often not used or is misused."

Districts must be prepared to invest considerable thought and effort in the assessment process. Even when using ready-made instruments developed by experts, users must exercise judgment in choosing appropriate instruments, administering them fairly and ethically, and then using the results in a productive way.

FORETASTE

Defining leadership is like defining love: you can put the words on paper, but somehow they never seem to quite capture the experience. Like love, leadership has many dimensions and has generated dozens of theories, and no single test provides a comprehensive measure.

Over the years, various researchers have located the roots of leadership in personality traits, behavior, the behavioral context, or personal inspiration. The contending viewpoints lead to instruments with very different content and format, none of which has emerged as an all-purpose tool.

In addition, potentially crucial issues such as values or racial and gender issues have not been extensively addressed by test publishers. Most important for K-12 administrators, measurement of school leadership (as opposed to corporate or political leadership) is rare.

The implication is clear: test users must understand the assumptions and definitions embedded in a leadership instrument before they can be sure it matches their needs.

This chapter's contents:

Looking at Leadership

Leadership Qualities

BEST COPY AVAILABLE

What Do Leadership Tests Measure?

Making a leadership assessment can be a deceptively simple process. Sitting at a desk for half an hour, an executive responds to a series of short questions, and in a short time (often just minutes) is rewarded with an impressively detailed profile showing his or her strengths as a leader.

But this seemingly straightforward process masks considerable ambiguity. Measuring leadership is not like measuring temperature, where we can expect that any well-made thermometer will give us pretty much the same result. Every leadership test makes different assumptions about what leadership is and how it can be recognized; every test uses its own language, format, and measurement strategies.

Thus, potential test users must understand how a given test defines leadership—what it includes and what it leaves out. This chapter provides a foundation for that understanding by examining the different conceptions of leadership and how they might be measured. The first half looks at the major theoretical perspectives that influence the content of leadership tests; the second part discusses some of the specific traits, skills, and attitudes that are typically measured on tests.

Looking at Leadership

Defining leadership is a lot like trying to dismantle a marshmallow: you can do it, after a fashion, but not very precisely, and not without getting your hands sticky.

Although people have always recognized leadership, only in the last two centuries has it been systematically studied (Bernard Bass).

In that relatively short time, hundreds of philosophers and researchers have dissected the concept, hoping to find some "essence" that could explain why certain people are admired, heeded, and followed by others.

As a result, there is no shortage of ideas on leadership. The most recent edition of the *Bass and Stogdil's Handbook of Leadership* runs to over 1,100 densely packed pages of research summaries. All this effort has not yet converged, and likely never will, on a single, universally accepted view. There are dozens of theories and hundreds of definitions, not to mention a seemingly infinite number of books and seminars offering "leadership secrets."

Although this situation is frustrating for school leaders who merely want some practical guidance, it should not lead to despair or cynicism. Even when theories differ, they are not always mutually exclusive; frequently they just reflect different dimensions of a very complex phenomenon. Any theory can be seen as a filter that highlights certain elements of leadership while ignoring others; collectively, multiple theories give us a rich vocabulary and a multilevel understanding.

However, this wealth of viewpoints creates a practical problem for those wishing to use leadership tests. In effect, every instrument offers its own operational definition of leadership. When a test asks leaders how often they delegate tasks, it asserts that delegation is an important dimension of leadership; when it asks how often leaders return phone calls, it assumes that communication is essential to leadership.

Test makers may or may not be conscious of any particular theory when they write tests, but every test reflects some implicit view of leadership. The following sections outline the major theoretical perspectives and how they are reflected in assessment.

Leadership as Personal Traits

When the earliest thinkers on leadership looked at greatness they saw unique, almost godlike, qualities. Thomas Carlyle, who believed that all of human history was shaped by "the Great Man," boiled it down to a simple formula: "The great heart, the clear deep-seeing eye: there it lies." Leadership came from personal qualities that rose above run-of-the-mill humanity.

Modern images are less heroic (and less gender-specific), but both researchers and the general public have frequently looked for the sources of leadership in biography. When we examine the lives of leaders, it is often difficult to avoid explaining their differences by personality, as when we contrast two twentieth-century politicians, Theodore Roosevelt and Robert A. Taft.

In Roosevelt we have a human dynamo whose high-energy approach drove him to leap at problems and throttle them into submission. Taft, an Ohio senator who was a leading presidential contender, took a much more passive stance:

> I don't try to supply great moral illumination or stir people up. There's probably only one man in a generation who can do that. What I do is work out each problem as well as I can, and then if my solution suits other people, they go along. (Joseph and Stewart Alsop).

Each politician's approach was so characteristic that it seems more plausible to explain the contrast by personality differences than by any conscious strategy. (It may also explain why Roosevelt became president and Taft did not.)

The first generation of leadership research assumed that leaders possessed distinctive personal characteristics that made them stand out from other people. These traits were regarded as stable, hence measurable through psychological testing. Researchers studied leaders in all walks of life, from corporate executives to athletes to recreation directors. Although they found a large number of traits associated with leadership, they were unable to formulate any simple explanation. Good leaders, it seemed, were intelligent, self-assured, enthusiastic, healthy, energetic, sociable—but not always. At best, certain traits might be necessary, but never sufficient.

As the list lengthened, without showing any signs of settling on a few key traits, researchers became discouraged and concluded that there was no "essence" of leadership that will hold for all cases. As Warren Bennis remarked of the leaders he studied, other than demonstrating certain similar abilities, they were "tremendously diverse. They were tall, short, fat, thin . . . . They evinced no common pattern of psychological makeup or background."

In recent years, researchers, reassessing the role of traits, have conceded that personal qualities do have an impact on leadership (Karin Klenke1996). While few scholars are ready to argue that certain people are "born to lead," it seems unreasonable to deny the

relevance of personal traits. The failure to reach consensus may simply reflect the great diversity of leaders who have been studied or the wide range of methods that have been used.

What kind of traits are we talking about? Klenke suggests three categories: physical (such as height, weight, and appearance); mental (such as intelligence); and personal (such as adaptability, self-confidence, and extroversion).

Typically, personal traits are measured by questionnaires given to leaders and their associates. Originally, psychologists sought comprehensive measures of broad traits such as intelligence and personality, but today's leadership tests have more modest aspirations, seeking only to zero in on selected personal characteristics that may influence leadership.

For example, *The Comprehensive Leader* (HRDQ) seeks to measure a trait called "Knowing Yourself," which reflects the leader's awareness of personal qualities. Likewise, one theme of *The Principal Perceiver* (Gallup Organization) is called "Achiever," defined as "an inner drive which continuously propels them to make things happen and get things done."

Personal traits can be assessed in a number of ways. The simplest approach is to list a number of statements and have leaders indicate the degree to which each statement is true of them:

- I am very generous
- I enjoy novelty
- I am uncomfortable with conflict

Similar statements can be given to superiors, subordinates, or peers of the leader to get a "360 degree" view. In scoring such tests, clusters of related statements are usually grouped into themes. For example, the *Life Styles Inventory* (Human Synergistics) groups its questions into a dozen categories, such as "affiliative," "perfectionistic," and "oppositional."

Another approach, used by the *Principal Perceiver*, asks open-ended questions and analyzes the leader's statements for key themes. This approach is less common because of the time involved for scoring and assessing answers.

Despite the somewhat diminished appeal of the trait approach among researchers, many practitioners continue to link leadership with personal qualities and thus find it meaningful to assess such qualities.

Leadership as Behavior

When researchers became disenchanted with the search for essential personal traits, they speculated that *behavior* rather than personality might separate successful from unsuccessful leaders.

Investigation showed that most leader behaviors could be grouped into two broad categories called "task orientation" and "relationship orientation." Task-oriented behaviors are aimed at getting the job done: communicating expectations, evaluating results, planning projects, and so forth. Relationship-oriented behaviors involve the interpersonal dimensions of leadership: conveying trust, empathizing, resolving conflicts, and related behaviors. Over the years these two categories have proved to be a useful way of describing leader behavior.

Although researchers immediately began to ask which set of behaviors was most important, they have never arrived at a clear answer. Hence, leadership instruments often attempt to measure both kinds of behavior, though some emphasize one more than the other.

For example, the *Educational Administrator Effectiveness Profile* (Human Synergistics) offers eleven themes with a strong slant toward task orientation: setting goals and objectives, planning, making decisions and solving problems, managing business and fiscal affairs, assessing progress, delegating responsibilities, communicating, building and maintaining relationships, demonstrating professional commitment, improving instruction, and developing staff.

Conversely, the *Principal Perceiver* has a strong relationship-orientation, including themes such as developer, relator, individualized perception, stimulator, and team.

In most cases, behavior is measured by having respondents indicate the degree to which they engage in particular behaviors. For example, the "communicating" cluster might include items such as:

- tells employees what is expected of them
- promptly communicates changes in plans
- provides feedback on employee performance

As with the measurement of traits, some instruments supplement the leader's self-assessment by asking associates to indicate how often they have observed these behaviors.

Another approach to measuring behavior is the assessment center. For instance, a leader may be given a simulated task (such as the well-known in-basket exercise) while trained observers assess his or

her performance. The National Association of Secondary School Principals runs assessment centers that examine behaviors in such areas as organizing, communicating, and problem-solving. The National Association of Elementary School Principals offers a "Professional Development Inventory" that asks administrators to respond to simulated leadership situations.

In assessment centers, the leader's behavior is directly observed by trained assessors. This method may give a more accurate picture than the informal impressions of leaders and their associates. However, centers offer a rather controlled "laboratory" environment and may not always detect how a leader *typically* behaves in a real-world setting.

Situational Leadership

When researchers first began exploring task-orientation and relationship-orientation, they thought that one or the other might prove to be more important to leadership. Despite dozens of studies, however, they failed to establish consistent links between particular behaviors and organizational outcomes. That is, some studies showed that task-oriented leaders were more successful in accomplishing goals while relationship-oriented leaders created better morale, but other studies showed that both types of behavior were associated with both kinds of success (Martin Chemers 1997). Once again, leadership recipes proved elusive, and some researchers speculated that leadership success depended on the situation; something that worked in one setting might fail in another.

This notion was intuitively appealing, since it might explain numerous cases where leaders of proved ability are unable to repeat their success when they transfer to new jobs. For example, a declining school with a demoralized, inexperienced faculty may benefit from a principal who provides clear goals, firm structure, and continual enthusiasm. That approach may be less successful at a thriving school with experienced, confident teachers who resent someone trying to tell them how to do their job.

In testing this idea, researchers have found a complex picture (Chemers). For example, one early finding was that task-oriented leaders are more successful in high-control situations (with a clear task, motivated followers, and high authority) *and* in low-control situations (with an unclear task, disgruntled followers, and limited authority). In the first case, a strong task orientation helps keep the

group happily on task; in the second case, it provides at least some order in a chaotic situation. By contrast, relationship-oriented leaders seem to do better with moderate control situations, such as motivated followers and an unclear task (where participative decision-making may be productive) or disgruntled followers and a clear task (where personal attention may sooth the feelings of followers).

The best-known application of this research comes from the work of Paul Hersey and Kenneth Blanchard (1993), who emphasize the importance of the match between the leader's behavior and followers' "developmental readiness." They describe two broad types of leader behaviors: directive and supportive. Leaders provide direction when they set goals, communicate expectations, and monitor results. They provide support by listening, encouraging, reassuring, and soliciting participation. Taken together, these two kinds of behavior lead to four leadership "styles:" high direction-low support (directing); high direction-high support (coaching); low direction-high support (supporting); and low direction-low support (delegating).

Followers can be characterized by their commitment and their competence. Taken together, these two qualities result in four different developmental levels that can be matched to the four leadership styles. Ideally, leaders will adapt their style to suit the need of the workers. Thus, employees with high competence and high commitment are best led through delegating; workers with high commitment but little competence respond well to directing. For Hersey and Blanchard, there is no "one best style"; the best approach depends on the situation, and the best leaders are flexible in their styles.

In theory, tests can assess this kind of flexibility by asking leaders to devise strategies to handle different kinds of situations. One test that does this in a direct way is *Leader Behavior Analysis II* (Blanchard Training and Development). Respondents are presented with twenty hypothetical scenarios and asked to choose the strategy (out of four listed) that would best represent the approach they would take. A companion instrument asks peers, superiors, and subordinates to predict how the leader would be likely to respond.

However, others find this approach to be a little too pat, suspecting that life is messier than a few hypothetical situations on paper. In addition, some researchers are skeptical of the Hersey and Blanchard model because its validity has not been extensively studied (Chemers). Few other tests attempt to measure situational appropriateness so directly.

However, many instruments *do* assume that leadership success depends on a good match between the leader's behavior and the demands of the situation. For example, tests of leadership "style" are based on the belief that leaders should at least be aware of their typical approach so they can recognize its advantages and disadvantages in different situations. Style generally refers to the characteristic ways that leaders make decisions, use power, and interact with others, but the specific definition can vary greatly.

For example, the Myers-Briggs Type Indicator (Consulting Psychologists Press) assumes that human personality is structured along four dimensions, each containing two contrasting possibilities: intraversion/extraversion, sensing/intuitive, thinking/feeling, and judging/ perceiving. These four possibilities result in sixteen basic "types," each with a distinctive flavor and a unique way of approaching life (Isabel Briggs Myers and Peter Myers 1980).

By contrast, the *Life Styles Inventory* (Human Synergistics) postulates twelve styles, each of which can be present in varying degrees: humanistic-encouraging, affiliative, approval, conventional, dependent, avoidance, oppositional, power, competitive, perfectionistic, achievement, and self-actualizing.

Despite such differences, most style measures seem to share several assumptions:

1. Style is a matter of relatively strong preferences that are difficult, but not impossible, to modify.

2. In general, no one style is preferable to others; each has advantages and disadvantages. For example, the *Life Styles Inventory* notes that those with a very strong "oppositional" style may be perceived as harshly critical; those with a very low oppositional style may be seen as weak or naive. Interpreters of the Myers-Briggs Type Indicator often point out the frustrations and ill feelings that can result when coworkers attempt to use clashing styles on the same problem (Robert Benfario 1991).

3. Style can be measured by assessing attitudes and perceptions as well as behaviors.

Transformational Leadership

Most of the viewpoints discussed above are leader-centered, focused on how leaders think and act. But leadership is also a social interaction between two or more people. (As many observers are fond of putting it, a leader is someone who has followers.) In a society that

values freedom and personal independence, what causes people to comply with the wishes of others?

Most researchers have explained this relationship as a form of *exchange*; that is, leaders ask for compliance from followers, but in return they provide concrete rewards such as money, recognition, and personal advancement. Seen this way, leadership is simply a rational economic transaction offering something for everyone.

In recent years, some observers have noted that the relationship sometimes transcends economics. Some leaders have followers who do more than comply or cooperate—they become enthusiastic and deeply committed, even when the leader has little material reward to offer in return. These *transformational* leaders build a sense of community, use evocative, emotionally charged language, and frame issues in moral terms (Bass).

Although this is sometimes called *charismatic* leadership, seemingly implying a return to Carlyle's brand of heroic leadership, the underlying issue is not personality but the way that leaders use authority. Whereas transactional leaders aim for *compliance* by offering concrete rewards to followers, transformational leaders aim for *commitment* by convincing followers that the organization is an extension of their identity.

In more concrete terms, transformational leaders use four tools to get results (Bernard Bass and Bruce Avolio 1994).

- *Individualized attention* recognizes the differences among followers and allows for their developmental needs.

- *Intellectual stimulation* turns the attention of followers to goals, aspirations, and new ways of doing things.

- *Inspirational motivation* is the way that transformational leaders help followers find meaning in their work.

- *Idealized influence* occurs when the leader serves as a living example and role model for followers.

Transformational leadership can be measured by tapping the perceptions of leaders and followers. For example, the *Leadership Practices Inventory* (Jossey-Bass/Pfeiffer) shows a transformational orientation when it seeks to measure categories such as challenging the process, inspiring a shared vision, enabling others to act, modeling the way, and encouraging the heart.

The Unknown Leader

If researchers have not yet completely put together the leadership puzzle, it may be because a few pieces are missing.

First, leadership theories—and leadership tests—are always based on the leaders we have. Most views of leadership were formulated in a time when the standard image of leaders was white and male. As women and minorities enter leadership roles in increasing numbers, "typical" leader behavior may change—along with our sense of the possibilities.

At the moment, there is no conclusive evidence that women or minorities develop distinctive styles that differ from traditional patterns (Chemers; Klenke). However, it may be that female and minority leaders have felt a need to "fit in" with prevailing leadership models; as their numbers increase, they may feel freer to develop new styles and strategies. Thus, test users should be alert to the possibility that current instruments do not fully reflect what leadership could be. For example, Klenke suggests that some tests have relied excessively on forced-choice opposites that limit response to a narrow range of possibilities.

Another area that has received limited attention is values. In recent years, leadership gurus have become fond of saying, "Leaders should spend less time worrying about doing things right and more time worrying about doing the right thing." Warren Bennis (1984), after a lifetime of studying leadership, says that the essential ingredients of leadership are our "best qualities": integrity, dedication, magnanimity, humility, openness, and creativity.

Although values have not been extensively studied, there is ample anecdotal evidence that followers respond well to qualities such as honesty and integrity. Kouzes and Posner, after surveying thousands of employees, have found that "honesty" is the main quality that followers want to see in leaders. In addition, studies point out that transformational leaders get their results by appealing to the values of their followers.

While some tests make an effort to measure core values such as honesty, integrity, and respect, none address a related question: Does the *content* of the leader's beliefs make a difference? After all, a transformational leader like Martin Luther King, Jr., is remembered not just for his qualities of dedication, integrity, and courage, but for *what* he stood for. If King had advocated a "go-slow" approach to civil rights, or had favored violent resistance, would he have accomplished as much?

The same question can be asked of school leaders. Thomas Sergiovanni has recently suggested that effective principals view learning in constructivist terms. Is this the case, or should principals follow the advice of the previous decade by providing highly structured, goal-oriented learning activities? Or does the answer vary from one situation to another? Research holds no answers, and leadership tests make no attempt to identify substantive values, much less suggest that a particular set of values is preferable.

Finally, no matter how eloquent or provocative the discussion or theory, empirical knowledge about leadership is limited by the lack of consensus on certain key questions, each of which is answered differently by different studies:

1. *Questions of definition.* Who is a leader? Is it someone who has formal authority? Someone who has technical competence? Someone who has persuasive ability? Is there a difference (as many claim) between "leaders" and "managers?" If so, how do we tell the difference? If a test says it has been validated on a population of managers, can we assume it will give us a portrait of leadership?

2. *Questions of effectiveness.* Is there a difference between successful and unsuccessful leaders? Studies of leadership have often built their conclusions around an identified population of leaders without asking whether these leaders were all successful. Leadership tests may also fail to make this distinction.

3. *Questions of context.* To what extent can results from a study be generalized to a larger population? Not all organizations are alike; what's required for success in one arena may be irrelevant in another. We can easily imagine, for example, that successful leadership in the microprocessor business requires boldness, imagination, and an ability to spot trends before they have fully emerged. Those same qualities are probably less useful in running a nuclear power plant. This is a significant question for school leaders, since many of the available tests have been developed in a corporate environment.

4. *Questions of bias.* Many leadership studies (and most leadership tests) are based not on controlled observation of behavior but on *perceptions* of behavior. Asked to rate their own behavior, leaders (like everyone else) may see themselves in a more positive light than actual observation would reveal (Bass). For this reason, leaders' perceptions are often cross-checked by tapping the perceptions of those who know them.

However, even fair-minded observers do not always see clearly. Chemers notes a persistent problem that comes from the "romance of

leadership"—the belief that organizational outcomes are always due to the actions of a heroic leader. When things are going well, or employees are happy, they are likely to attribute it to the actions of the leader. As a result, when surveyed, they may "see" positive behaviors that have not actually occurred. In short, leaders' actual behavior is obscured by the glare from their haloes. (Of course, if the organization is going through hard times, the reverse may apply.)

These questions do not invalidate existing leadership studies or the instruments that result from them. But they do suggest that tests of leadership should always be considered provisional or tentative.

The Nature of School Leadership

The views of leadership discussed in the preceding section have originated in the corporate world (or, to a lesser degree, in politics). Can we safely assume that school leadership follows the same rules?

Educational administrators have often been quick to take their cues from business. For example, Total Quality Management—a process designed to reduce manufacturing defects—took just a year or two to emerge in education as the Total Quality School.

Yet the differences between business and education are significant:

- Schools are not profit-oriented, so there is no single standard of success (test scores are highly ambiguous).
- School leaders operate in a highly public domain, knowing that every action will be scrutinized, and every memo may become a public record.
- Schools are political institutions trying to please an increasingly diverse and contentious society; consensus—or even the appearance of consensus—is difficult to achieve.
- Schools have a near monopoly, meaning the competitive spirit that drives many businesses is lacking.

Although school superintendents are often compared to CEOs, it's hard to imagine a private-sector executive with a similar task: managing a well-educated group of professionals while simultaneously cajoling and soothing a lay board and ever more demanding public.

Thomas Sergiovanni (1996) argues that schools are an inherently *moral* enterprise requiring a very different approach than that used by

business. "Wholesale borrowing of theories from other fields has resulted in a characterless literature of educational administration," he writes.

> Our quest for character does not begin with some other field's models, but with our own visions of what we want for the parents and students we serve, of how we want to work together to achieve these purposes, and of what we need to do for purposes and beliefs to be embodied in school practice.

Sergiovanni argues that an adequate theory of educational leadership should be aesthetically pleasing, using language that is both appropriate and appealing. (For example, gardening is a more appropriate metaphor for teaching than Total Quality Management.) It should place moral authority at the center, emphasizing the mutual connections and obligations that link members of the school community. It should emphasize the school as a producer, not just a transmitter, of knowledge.

In practical terms, this view demands that principals have skills in nine areas:

- purposing (transforming shared visions into a moral covenant)
- maintaining harmony
- institutionalizing values
- motivating
- managing
- explaining
- enabling (removing obstacles to participation)
- modeling
- supervising

While some of these skills can be found in various leadership instruments, no test has yet been built around them or, for that matter, around any comprehensive theory of school leadership.

William Greenfield (1986) advances a similar view, noting that schools are *normative* institutions, gaining employee commitment through persuasion and appeals to common values rather than coercion. Principals have some power to reward and punish, but teachers who are so inclined can easily evade the wishes of principals.

Greenfield suggests that leadership in this setting requires five sets of skills:

- Managerial (coordinating, directing, and monitoring school operations)
- Instructional (overseeing the school's core activities of teaching and learning)
- Social (face-to-face interactions with students, teachers, parents, and community members)
- Political (negotiating the inevitable clashes among the multiple agendas of members of the school community
- Moral (determining what policies and practices are desirable and consistent with the school's purpose)

Finally, there are those who see school leadership as a *craft*. Arthur Blumberg (1989) defines *craft* as "the idiosyncratic use of self to make prudent decisions concerning problematic situations in school life." Craft includes standard administrative skills, but also involves a comprehensive, detailed knowledge of a particular school—what Blumberg calls "a nose for things." He quotes one principal who seemed to exemplify "management by walking around":

> I can sense when things are going right or wrong on any particular day. When I walk around, I listen for the level of noise in the building and I look for the amount of debris or litter. If it seems more than normal, it's a signal that some vandalism may be creeping in.

How does one assess this skill? While a leadership test could tell rather easily whether principals spend time in the hallways, it would be much more difficult to determine whether they were managing or just walking around. Administrative craft seems to involve a great deal of *local* knowledge of a kind that would be difficult for a national test to assess.

In general, the measurement of school leadership (as opposed to other kinds of leadership) is relatively rare. Most tests come from the corporate world, and while they can offer excellent insights into school leadership, they should be used with care.

More generally, however, we can say that *no* test (even one designed specifically for school environments) will offer a comprehensive picture of leadership capability. For now, at least, school officials who are looking for a simple off-the-shelf solution to leadership assessment will be disappointed.

Leadership Qualities

The first part of this chapter examined some theoretical views of leadership and how they are reflected in leadership tests. This section looks more concretely at some of the specific traits, behaviors, and attitudes that frequently show up on leadership instruments. This discussion is not all-inclusive, but demonstrates in practical ways the form that these qualities take on tests.

Communicating

Virtually everyone agrees that communication, though it is defined in a variety of ways, is essential for leaders. Sometimes communication refers to the leader's ability to provide information and directions in a timely way; other times it involves the ability to establish and maintain mutual dialogue. Some sample test items might look like this:

- I never end a meeting until I make sure that everyone understands the decision.
- Before any decision is made, I solicit opinions from those who will be affected.
- I maintain an open-door policy.
- I make an effort to return all phone calls on the day I receive them.
- I frequently walk around the building so I can interact with people I might not otherwise see

Establishing Relationships

Leadership is a people-intensive business, and many tests attempt to find out how effectively leaders establish and maintain personal relationships. Some sample items:

- I take a personal interest in my employees.
- I make sure that employee achievements are recognized.
- I enjoy being around people.
- I frequently chat with people about things that have nothing to do with work.

- When I have a conflict with others, I always listen to their point of view before I explain my position.

Goal-Setting and Planning

Most people believe a major leader role is providing the organization with a clear sense of direction and taking the steps to ensure that the destination is reached. Some sample items:

- At the beginning of each new project I make sure that everyone understands the goals.
- I make sure that each goal we set will get the necessary resources to carry it out.
- I frequently talk about our goals so that everyone understands their importance.
- At the end of each year I update our goals to ensure they are still relevant.
- I establish careful timelines for each goal.

Empowering Others

Recent views of leadership have emphasized the leader's role as a facilitator: someone who leads through others. Some sample items:

- I allow others to work things out for themselves even when I could do it more quickly myself.
- I provide resources and encouragement for employees to upgrade their skills.
- I frequently ask staff members for their views.
- At meetings I make a deliberate effort to sit back and listen rather than leading the discussion.
- Once I establish a committee to accomplish a task, I let them do their work.

Inspiring

The recent emphasis on transformational leadership has created an interest in leaders who use their moral authority to inspire and motivate others. Some possible items:

- I frequently talk about our vision for the future.
- I like to tell stories about people who have really lived up to the ideals of this organization.
- I make an effort to hire people whose values are aligned with our mission.
- I demand that employees live up to the ideals of this organization.
- I make sure we celebrate the accomplishments that lead us closer to our vision.

Ethics

James Kouzes and Barry Posner (1993) note that the leadership quality most often valued by subordinates is honesty. Leaders who conduct themselves ethically, recognizing and acting on their responsibility to others, will gain credibility and moral authority. Some sample items:

- I am willing to admit it when I've made a mistake.
- When I've given my word, people know I'll follow through.
- I don't ask my subordinates to do anything that I'm not willing to do myself.
- I'm willing to take public criticism in order to uphold our mission.
- I practice what I preach.

Style

Style is a broad concept that can include many of the behaviors and attitudes already discussed. Most often, however, it focuses on the characteristic ways that people solve problems and interact with the world around them.

- I would rather deal in concrete data than in theorizing about possibilities that might never occur.
- When faced with a difficult decision, I listen to my heart rather than my head.
- Interacting with large numbers of people keeps me energized and interested.

- I would rather keep my independence than be indebted to others.
- I'd rather complete one task at a time instead of juggling several at once.

Personal Traits

Although the trait explanation of leadership has lost favor, theoreticians and test makers still see certain personal qualities as playing a significant role. For example:

- When things get busy, I'm able to find the extra energy to keep going.
- I don't lose my temper or take out my frustrations on others.
- I take a lot of work home with me in the evening or over weekends.
- I'm always reading the latest research to find new ideas I can use.
- I'm very hard on myself.

These examples give the flavor of many paper-and-pencil assessments, and also reveal some of their advantages and disadvantages. The items are usually simple and can be answered without a great deal of analysis, particularly when they focus on concrete behaviors. A fairly compelling portrait can be created in a short time.

However, such items can be highly subjective; words such as "frequently," "usually," or "a lot" leave considerable room for interpretation. In addition, items of this type do not provide reliable measurement of actual skill or effectiveness. Thus, results will be most useful when used as part of a comprehensive process of analysis, reflection, and goal-setting. (See chapter 4.)

A somewhat more direct measurement of leadership capability comes from asking leaders to respond to simulated problems. For example, an inbasket exercise might provide the following data:

> Sitting at your desk at 7:30 Tuesday morning, you find the following items on your to-do list:
>
> 1. Prepare an agenda for the curriculum committee meeting on Thursday.

2. Check on a report of increasing student rowdiness at bus unloading.

3. Meet with Sylvia LeMahieu, a third-grade teacher who has frequently arrived late to school this year.

4. Consult with architect about this summer's renovation project.

5. Prepare for meeting with unit heads to discuss a mandated budget reduction of 10 percent.

6. Messages on your answering machine:

—Arlene Johanson, PTA president, wants to discuss the reading carnival scheduled in two weeks.

—John Barrett, the superintendent, wants you to call.

—Ed Arnett, a reporter for the local paper, wants a comment on your school's recent decline in test scores.

—Myrt Halverson, owner of the bus company, wants to discuss a possible route change.

—Mrs. Cantor, parent of a sixth grader, is upset about the amount of homework he has been receiving.

—Leann Alvah, candy company representative, wants to discuss a possible fund-raising opportunity.

In responding to this agenda, leaders will reveal a great deal about their thought processes and their ability to prioritize and resolve typical administrative problems.

Alternatively, leaders might be given a hypothetical scenario and asked to explain what they would do and why:

In the two years that you have been principal of Glen River High School, you have received a growing number of complaints about Ed Halley, a veteran English teacher. The complaints are always the same: poor discipline, uninspired lessons, and a failure to challenge students.

You know that Halley was once considered a competent, dedicated teacher. Indeed, his dedication has never been in doubt; he often arrives at school early, is seldom absent, and

has often been a moderating voice among more militant teachers. Your observations of Halley's classroom have confirmed many of the complaints. Halley often seems to be going through the motions; his lessons are flat and lifeless, consisting mostly of teacher monologue. While student behavior is not extreme, there is far too much restlessness, inattention, and chatter to permit much learning. Halley seems to ignore the lack of student attention; when he does admonish students, they feel free to ignore him. On at least two occasions, you have seen students talk him into canceling a homework assignment.

You have raised these issues in postevaluation conferences, but Halley denies there is a serious problem. He concedes that students are not wholly attentive, but blames it on television and lack of discipline at home: "I don't like it either," he says, "but you can't fight a whole society." At 55, he admits he "may have lost a step or two," but pointedly adds, "Aging isn't a crime."

1. In resolving this dilemma, what are the key factors you would need to consider?

2. What additional information would you need before making a decision?

3. Based on the information provided, what action would you take?

Items such as these provide good insight into a leader's ability to recognize critical elements in a problem and to make a well-reasoned decision. Obviously, there is no single "right" answer, but test takers would be expected to address the key issues. While this type of question provides information not available through simple self-inventories, it also requires considerably more time to take and to score, and the need for trained assessors makes such tests rather expensive.

The Need to Choose Wisely

As this chapter has demonstrated, experts are far from a consensus on the nature and demands of leadership, and leadership assessments reflect these differences of opinion. The diversity of test instruments provides schools with both danger and opportunity. The

danger is assuming that any test of leadership will be suitable for a particular school; the opportunity is that in the wide range of choices, school officials are likely to find a test that suits their specific needs. Chapter 3 discusses the steps required to make a wise choice.

FORETASTE

This Chapter's Contents:

Step 1: Know Your Purpose

Step 2: Develop a Local Definition of Leadership

Step 3: Locate Possible Instruments

Step 4: Determine the Suitability of Various Tests for Your Needs

Step 5: Establish the Credibility of the Instrument

Step 6: Consider Practical Issues

Step 7: Consider the Ethical and Legal Issues

Step 8: Determine How Much Support the Instrument Provides for Followup and Professional Development

Selecting the Best Instrument

If the world of leadership theory is multifaceted and inconsistent, so is the world of leadership testing. Organizations can choose from dozens, perhaps hundreds, of instruments designed to help select or develop leaders.

School districts looking for "the one right test" are likely to be disappointed. No instrument is perfect (especially since relatively few have been specifically designed for school settings), yet more than one may be acceptable. Making the right choice requires careful attention not only to the instrument, but to the needs of the district.

Thus, buying a test off the shelf because it says "leadership" is a bit like buying a mail-order suit because it has three buttons and a neat little pinstripe: you'll get what you paid for, but it won't necessarily fit. The consequences of a poor choice can be significant, ranging from wasted time and money to poor personnel decisions. This chapter outlines a decision-making process that increases the chances of an appropriate choice.

Step 1: Know Your Purpose

Schools may want to measure leadership for several distinct purposes, each of which will raise a particular set of issues.

Selection. As noted in chapter 1, leadership tests can enhance traditional methods of selecting leaders, such as interviews. However, using tests for selection brings some special responsibilities. Hiring is a high-stakes decision; districts may incur considerable expense and aggravation from a poor choice. Using tests for selection requires careful job analysis to determine what qualities and skills are desired for a given position. Candidates have a lot at stake as well,

and the possibility of discrimination raises both ethical and legal issues. *Few of the instruments discussed in this book have been designed for selection purposes.*

Evaluation. Using leadership tests to evaluate performance is highly controversial, especially when compensation or career advancement will be affected (Maxine Dalton 1998). At best, tests are an indirect measure of leadership, always one step removed from the arena of action. Ideally, leaders who have been employed by a district for any length of time should be rated by on-the-job performance, not by the refracted images that come from tests.

Dalton notes that linking test results with concrete rewards violates the conditions of "psychological safety" needed for productive feedback. The greater the stress, the less likely that leaders will be able to accept and learn from feedback that clashes with their self-image. In addition, tests relying on feedback from peers and subordinates may produce inaccurate results because of reluctance to do harm to colleagues.

Ellen Van Velsor (1998) suggests that in organizations with a well-established tradition of using tests for self-development, where employees are comfortable and experienced with the feedback process, it may be possible to use tests for appraisal. However, she suggests that when testing is being introduced for the first time, linking it to formal evaluation would be a mistake.

At the same time, leadership instruments may appropriately play a role in the overall assessment process. Administrator evaluation often includes self-assessment and goal-setting, and the right instrument can enrich the reflective process. To do this effectively, districts must be sure that their evaluation process establishes a link between on-the-job performance and the qualities measured by the instrument.

At a minimum, districts initiating a feedback process should use great clarity in explaining the relationship between the testing process and formal appraisal. In particular, leaders should know what limits, if any, are being placed on anonymity and confidentiality (Van Velsor). *None of the instruments discussed in this book have been specifically designed for on-the-job evaluation.*

Professional Development. Most often, leadership tests are used to stimulate professional development. A well-chosen instrument may give leaders new insights into their own behaviors, making them conscious of strengths and weaknesses. In such cases, schools must ensure that the chosen instrument is designed to promote reflection and provides followup and support for the leader's self-analysis.

Most of the instruments described in this book are designed for professional development.

Step 2: Develop a Local Definition of Leadership

As chapter 2 made clear, definitions of leadership cover a wide range of personal qualities and skills, and each leadership instrument reflects a particular set of assumptions. To make an appropriate match, districts must have a clear conception of how *they* define leadership at the local level. That is, what particular qualities and skills are most important in *this* district at *this* time?

In many cases, districts have not given much explicit thought to defining leadership, but the process does not require starting from scratch. Helpful clues can be found in the leadership standards developed by professional associations as well as within the district itself.

A good starting point is to consider professional consensus on leadership essentials. In recent years, groups such as the National Policy Board for Educational Administration, the National Association of Secondary School Principals, the National Association of Elementary School Principals, and the American Association of School Administrators have developed standards to guide the preparation of school leaders, and they have generated lists of key school-leadership traits.

In 1997, the Interstate School Leaders Licensure Consortium, a collaborative effort among school leadership organizations and a number of states, released a comprehensive set of standards that will be used to guide administrator-preparation programs. The standards say, "A school administrator is an educational leader who promotes the success of all students by" doing the following:

- facilitating the development, articulation, implementation, and stewardship of a vision of learning that is shared and supported by the school community
- advocating, nurturing, and sustaining a school culture and instructional program conducive to student learning and staff instructional growth
- ensuring management of the organization, operations, and resources, for a safe, efficient, and effective learning environment

- collaborating with families and community members, responding to diverse community interests and needs, and mobilizing community resources
- acting with integrity, fairness, and in an ethical manner
- understanding, responding to, and influencing the larger political, social, economic, legal, and cultural context

The standards are accompanied by specific indicators. For example, the first standard includes knowledge indicators such as understanding "the principles of developing and implementing strategic plans," "systems theory," and "learning goals in a pluralistic society." Leaders should also be committed to such things as "the educability of all," "continuous school improvement," and "a willingness to continuously examine one's own assumptions, beliefs, and practices." Finally, the standards must result in actions, such as these pertaining to the implementation of a school's vision:

- Progress toward the vision and mission is communicated to all stakeholders.
- The core beliefs of the school vision are modeled for all stakeholders.
- Barriers to achieving the vision are identified, clarified, and addressed.

Collectively, these standards provide a comprehensive and credible foundation for leadership assessment. In fact, they provide the basis for a test developed by Educational Testing Service that is already being used in several states as a licensure requirement for new principals.

Districts doing their own assessments can draw on whichever sections of the standards seem most relevant to their current needs. Determining those needs takes some effort, but clues can be found in several sources.

1. *Job descriptions.* Most districts have some sort of job descriptions for leadership positions. Although these are often rather generic and uninformative, they can be transformed into focused questions about current leadership needs. For example, one of the described duties for a principal may be "to coordinate instructional improvement efforts." Given the school's history and current status, what sort of leadership is required? Are teachers experienced and confident, needing a facilitator more than a director, or do they require forceful leadership to set a direction? What other skills and qualities are important in carrying out this responsibility?

2. *Analysis of leadership history in the district.* Every district can recall examples of leadership that failed (perhaps spectacularly) as well as leadership that was highly successful. What was behind the failures? Could they be attributed to the lack of certain key characteristics? What qualities, skills, or attitudes do the successful leaders seem to share?

3. *Analysis of local conditions.* Where are the current controversies, sore spots, and hot buttons in the district? What are the district's most urgent needs? What kinds of leaders are most likely to maneuver their way through the current situation? For example, if relations with teachers or parents are strained, districts may want to give special attention to interpersonal skills.

4. *Vision.* At times, districts may not want someone who "fits in" with the status quo; instead they want someone who can take the school in dramatic new directions. If the district or school has a clearly stated vision, this will often suggest the kind of leadership required. An added benefit is that assessing leaders on these visionary qualities will send an important message that the district is serious about them.

5. *Individual self-analysis.* Even when a districtwide consensus on the nature of leadership already exists, each administrator develops his or her skills in a highly individualized manner. Every leader, at a given point in time, has a unique collection of technical skills and personal qualities—strong in some areas, weak in some areas, and nonexistent in still others. Each leader may need to know something different that a one-size-fits-all approach cannot provide. For example, if a leader is experiencing frustration with interpersonal conflict, an instrument that casts light on that issue would be especially useful.

Most leadership instruments are based on the assumption that leaders are autonomous professionals who must play an active role in the assessment process. Their involvement should be solicited at the beginning, not after an instrument has been chosen. Table 1 displays a series of questions that will draw out some of the issues that should be included on an assessment.

Step 3: Locate Possible Instruments

At one time, leadership instruments were difficult to locate and examine, but several sources now offer centralized listings, either published or online. Most vendors will also provide examination

TABLE 1

Leadership Self-Analysis

1. What leadership tasks have provided the greatest challenge for you in your current position? Why?
2. What skills have been the most useful to you in your current position?
3. What personal qualities have been most useful to you in your current position?
4. If you were hiring a replacement for your position, what skills and/or qualities would you look for?
5. What skills could you develop that would help you better meet the demands of your current job?
6. What skills could you develop that would best help your long-term development as a leader?
7. What would you most like to know about yourself as a leader?

copies or detailed descriptions of their tests. The instruments described in chapter 5 can provide a starting point, but prospective test users should check other sources as well.

One of the most convenient sources is the test-locator website maintained by the ERIC Clearinghouse on Assessment and Evaluation (http://ericae.net). The database provides quick access to information on more than 10,000 instruments of various types, including many on leadership. For each test there is a brief descriptive abstract and the publisher's address. In addition, the site can locate published reviews of some tests, and it also has a section providing guidelines for choosing instruments.

Another source is the leadership-education source book published by the Center for Creative Leadership (Frank Freeman and colleagues 1996). The book contains brief descriptions of seventy-five commonly used leadership tests, including information on target

audiences, administration procedures, and comments on validity and reliability. Vendor addresses are provided.

Information on the assessments conducted by the American Association of School Administrators, the National Association of Secondary School Principals, and the National Association of Elementary School Principals can be found by contacting those organizations.

Step 4: Determine the Suitability of Various Tests for Your Needs

The synoptic descriptions found in the literature can screen out many tests that are clearly inappropriate, but the final choice depends on a close examination of the instruments and supporting information provided by vendors. Three questions are particularly important:

1. *What are the intended uses of the instrument?* Test-makers design their instruments to fill a need, and the content, interpretation, and follow-through activities are tailored to fit that purpose. Most of today's commercially available leadership tests are designed for professional development, which means that their use for selection and evaluation is at least questionable, if not inappropriate.

2. *Who is the target audience?* Tests are usually designed with a particular audience in mind, and scoring is often based on comparisons within that target group. Most leadership instruments have been developed in and for a corporate environment. This does not automatically invalidate their use in school settings, since some leadership qualities undoubtedly cross occupational lines. However, even seemingly "generic" leadership qualities can differ from one setting to another. For example, communication with peers and subordinates in a corporate environment may be very different from communication with teachers and parents in a school setting. Because of the scarcity of education-specific instruments, this issue should be carefully considered.

3. *Does the content of the test address the leadership issues of concern to the district?* This is the essence of the analysis, one that can be approached through a two-part process. The first step is to develop a table that specifies the content requirements and then to check off which instruments address those issues (see table 2).

Second, the most promising instruments should be read to see whether the questions are actually congruent with the district's definitions. A term like "communication" covers a lot of ground; the

district may be interested in interpersonal relationships, while the instrument focuses on the dissemination of information.

Step 5: Establish the Credibility of the Instrument

Nothing undermines assessment faster than the perception that the process is inaccurate, irrelevant, or unfair. Anyone contemplating the use of a particular instrument should have solid evidence that the instrument is statistically sound and capable of providing the information desired—and that it will be seen that way by participants.

The standard statistical measures of soundness are reliability and validity. While test users will seldom need (or be able) to get involved with sophisticated calculations, they should be able to evaluate the information provided by test publishers about test characteristics.

Validity

Validity is the bedrock of a test's soundness, offering assurance that the instrument can deliver on its promises (American Educational Research Association and others 1985). Statisticians have developed numerous ways of testing validity, but two common-sense questions will get to the heart of things:

1. Does the instrument assess the qualities it claims to? Would a reasonable person agree that the items in this assessment reflect the particular dimensions of leadership that we're interested in? If the answer is "yes," the assessment will be seen as credible.

2. Is there evidence that performance on the assessment correlates with performance on the job? That is, do high scorers perform better on the job than low scorers? Not every instrument has been validated in this way, and even when the information is provided it has to be interpreted carefully. For example, results of tests developed in a corporate environment are not automatically transferable to school settings. Over time, however, you should be able to determine that a particular instrument is a good predictor of leadership success in your district.

These questions can be answered in a variety of ways. In answering the first question (Does the instrument measure the qualities it claims to?), we can attempt to gauge the instrument's face validity, content validity, and construct validity, then seek to determine its predictive validity.

TABLE 2

Sample Checklist To Match Tests with Desired Leadership Characteristics

	Vision	Communication	Empowerment	Risk-taking	Flexibility
Test 1	X	X		X	
Test 2		X	X		X
Test 3			X		
Test 4	X		X	X	X

a. *Face validity*. Face validity does not involve any technical tests; it simply means that a reasonably well-informed person would agree that the test *appears* to measure leadership. Obviously, this kind of impressionistic judgment offers no objective assurances, but if the answer is "yes," the test is likely to have credibility among those who use it.

b. *Content validity*. Like face validity, content validity reflects a judgment about how well a test covers a particular domain. However, the judgment is more formal; for example, test-makers may ask leadership experts to indicate how well the test measures essential leadership qualities.

c. *Construct validity*. A construct is a theoretical quality that is assumed to exist but cannot be directly measured. Leadership is a prime example. We observe differences in the ways that leaders act or in the results they get, and we infer that there must be some quality that accounts for it.

A test with good construct validity is one that is in accord with a particular theory of leadership. For example, the *Leadership Practices Inventory* (Jossey-Bass/Pfeiffer) seeks to measure exemplary leadership rather than routine management skills; the construct is centered on the way that ordinary people accomplish extraordinary things. Research by the instrument's authors found that the essential qualities of exemplary leaders were covered by the categories "challenging the process," "inspiring a shared vision," "enabling others to act," "modeling the way," and "encouraging the heart." The *Leadership Practices Inventory* would have good construct validity to the degree that the test items reflected these categories.

Note, however, that construct validity is not universal across instruments. Another test may have low construct validity in relation to the theory underlying the *Leadership Practices Inventory*, but have high construct validity for a different leadership theory. In general, tests are considered sounder if they show construct validity in relation to *some* theory; this indicates that the results are likely to be coherent and understandable, rather than a random collection of scores. For districts that have a well-articulated leadership theory of their own, the best instrument would be one that has high construct validity for that theory.

d. *Predictive validity*. While the above types of validity offer some assurance of credibility, most people are even more impressed by evidence that the instrument works in the real world. If high scorers on the test turn out to be effective leaders in everyday terms,

then we can use the instrument to predict success rather than simply waiting to see what happens.

Predictive validity is most often measured by correlating the test to some kind of external job-related criteria. For example, in the business world predictive validity is sometimes measured by administering a leadership test to a group of would-be executives and then returning in five years to see how high they have risen in the organization. Or, we could survey their subordinates to learn how they are perceived as leaders. (We could also do similar surveys of their current performance as leaders—a measure known as *concurrent validity*.)

While studies of predictive and concurrent validity yield hard numbers, those numbers are only as good as the external criterion being used. We might speculate, for example, that promotions are sometimes influenced by political connections or sheer luck. Surveys of subordinate perceptions may offer only a limited picture of leadership effectiveness. Thus, no matter how strong the validity scores provided by publishers, users must still make the final judgment that a test reflects the desired traits. As the authors of the *Standards for Educational and Psychological Testing* point out, "It is incorrect to use the unqualified phrase 'the validity of the test.' No test is valid for all purposes or in all situations" (American Educational Research Association and others 1985).

Reliability

The other statistical check of an instrument's soundness is reliability, which simply measures the likelihood that a test score is typical of the person achieving it. That is, would he or she achieve a similar score if the test were taken again? If not—if scores were to fluctuate dramatically from one time to the next—the results would be of little value.

Reliability can suffer for a number of reasons. The most obvious (and the hardest to control) are the condition and attitude of the test-taker. Someone who is mentally fatigued, depressed, distracted, or unmotivated is likely to produce a score that fails to reflect his or her true capabilities. Test publishers have no way of measuring the degree to which this has occurred in a particular case. It is best controlled at the time of the test, by establishing conditions that allow subjects to perform at their best. Motivation is especially important; if

test-takers see the test as irrelevant or inappropriate, they may not give it their full attention.

One way to raise reliability is to make the test long enough that it asks multiple questions on the same topic. A single question may be answered inaccurately for any number of reasons, such as momentary inattention, ambiguity, or stress; when a test asks multiple questions on a particular topic, the chance of an accurate answer rises.

Subjectivity also threatens reliability. Tests that require scorers to use personal judgment in evaluating the answers may result in inconsistent scoring. For this reason, tests requiring subjective judgments usually provide training or even certification for those who are administering them.

Reliability is measured in several ways. First, *split-half* reliability randomly divides the questions into two groups and checks to see that scores for each half are roughly similar. Second, in tests requiring subjective judgment, the scoring of different raters can be compared to see that they do not seriously diverge. Reliability scores range from 0 to 1.0; the closer to 1.0, the higher the reliability (scores above 0.8 are considered quite good).

Finally, *test-retest* reliability examines the stability of scores from one test administration to another. This is measured by giving a test to a group of leaders, waiting a month or two without providing any feedback, and then administering the test a second time to the same group. If scores differ greatly from one time to the next, the instrument's reliability is suspect. Usually, a reliability coefficient of at least 0.4 is acceptable, though correlations of 0.6 or 0.7 are more desirable (Van Velsor).

Reliability and validity scores are always less than perfect, but test users can add credibility by exercising care in the way they administer the tests:

1. *Minimize possible sources of unreliability.* Tests should be administered at a time when test takers are alert, unhurried, and protected from distractions. They should understand the purpose of the instrument and should agree that the test is credible for that purpose (or at least be willing to keep an open mind).

2. *Choose instruments that are well written and sharply focused.* Van Velsor points out that poorly constructed tests can affect the validity and reliability of the results:

- Vaguely defined words (such as "conservative" or "communication") can evoke vastly different responses from participants.

- Items with too few alternatives ("agree/disagree") can disguise important gradations of behavior; items with too many responses ("strongly agree," "agree," "slightly agree") can become meaningless.
- In using 360-degree feedback, items requiring a high degree of inference are likely to be viewed more skeptically. For example, items focusing on attitudes rather than observable behavior usually involve more guesswork. In addition, different groups of raters may not be equally positioned to comment on a leader's behavior. For instance, teachers are unlikely to know how a principal interacts with superiors.
- Multipart items lead to ambiguity. When an item says, "articulates a clear set of expectations and adheres to them," it is quite possible that the answer is "yes" to the first part and "no" to the second part.

3. *Ensure congruence between test content and the district's beliefs about leadership.* No matter how impressive a test's statistical validity, it will suffer a loss of credibility if it seems unrelated to the daily demands of leadership. When districts preach teamwork and collaboration but use instruments that ignore those dimensions, test takers will have little motivation to take the process seriously. When participants see little of their world reflected in the instrument, they will quietly reject its relevance.

4. *Determine that feedback scales are meaningful.* Most instruments group a large number of items into a smaller set of leadership characteristics. For example, five items may be grouped under "communication," five under "vision," and so forth. Ideally, each item in a particular category should be strongly correlated with the others and should clearly reflect the underlying "theme" of the category. For example, items such as "frequently articulates reasons in writing," "makes sure that others understand the rationale," and "frequently refers to established goals" seem to reflect "communication" and would probably be highly correlated. An item such as "listens sensitively to others' concerns" also involves a kind of communication skill but may not necessarily be found with the first three items. A coefficient of at least 0.7 is needed to have confidence in the reliability of the feedback scales (Van Velsor).

5. *Use 360-degree feedback.* An instrument's validity and reliability can be enhanced by including multiple perspectives in the assessment process. A growing number of tests supplement the

leader's self-assessment with feedback from supervisors, teachers, and peers. Each group sees leadership from its own limited and possibly biased viewpoint, but using all perspectives permits a reasonable "triangulation" of the leader's actual performance.

6. *Debrief participants after the test has been administered.* The ultimate credibility of the process will rest less with statistical analysis and more with the perceptions of participants. Their reactions will provide valuable clues about the meaningfulness of the instrument. The followup process (see chapter 4) should always leave room for the professional judgment of those who took the test.

Step 6: Consider Practical Issues

How long does the test take? Can it be scored locally? What's the cost? No instrument will satisfy all your purposes, and the final choice will be the result of numerous tradeoffs.

Step 7: Consider the Ethical and Legal Issues

Leadership instruments can have major consequences for those who take them. This is most evident when the tests are used for selection or evaluation, but even when limited to professional development they require the expenditure of valuable time and energy, and they may subject test-takers to disturbing feedback about their leadership skills. Districts have a major responsibility to ensure that their use of leadership instruments is both legal and ethical.

The Joint Committee on Testing Practices, a cooperative group initiated by the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education, has developed a "Code of Fair Testing Practices in Education" (1988). It identifies four major responsibilities for those using tests.

1. The user should thoroughly examine test content, manuals, and supplementary material to ensure that tests satisfy the purpose for which they were designed and that they are being used with appropriate populations.

2. Users should make sure that test results are interpreted correctly and used only for the indicated purposes. They should also

gather evidence to show that the instrument is fulfilling its stated purpose.

3. Users should select and use tests that have been developed to avoid bias related to sex, race, and handicapping position. Overt bias is less common today than several decades ago, but if results show differential performance related to gender or race, users should make an effort to determine why.

4. Test takers should be provided full information on the nature of the test and its intended uses. It is particularly important for them to know whether (or to what extent) their responses will be confidential.

Step 8: Determine How Much Support the Instrument Provides for Followup and Professional Development

Most formal assessments come with posttest activities that help leaders reflect on the results and set goals for continuing development, but they vary in quality. The key questions: Are the test results reported in a clear, understandable format? Do the supporting materials provide test-takers with insights into the meaning of their leadership profile? Does the test package help administrators set an agenda for improvement? Is there evidence that leaders who go through the testing and followup activities actually improve their skills? These issues come into greater focus in chapter 4.

FORETASTE

In many districts, assessment is an event; the feedback is delivered, the accountability report is filed, and everyone moves on to more pressing business. By contrast, best-practice assessment is a *process* in which providing feedback is just the first step in a much longer journey.

Districts wishing to get their money's worth from assessment must make sure that participants develop an accurate understanding of the feedback, that they use it as a tool to reflect on their strengths and weaknesses, and that they transform it into a plan of action.

This chapter's contents:

Selecting the Facilitator

Deciding on the Setting

Delivering the Feedback

Forming a Development Plan

Searching for Ground Truth

Fulfilling the Promise of Assessment

As she entered the superintendent's office, Anna Sebring found herself confronting mixed emotions about the meeting to come. As a first-year principal, she was eager to get some feedback on her performance, and the district's new 360-degree feedback program would give her information on how her teachers, peers, and superintendent perceived her leadership, as well as insights into her self-assessment.

At the same time, she was worried that the results might not show her in a favorable light. Things had gone fairly smoothly, but it was always hard to tell how others might feel. At times she even pondered the unthinkable: a message that she wasn't really cut out for administration.

She wished she had the easy-going confidence of Randy Armani, her veteran colleague from Everett School. "These things are all alike," he had told her with assurance. "They're looking for happy talk. So just go in, nod your head whenever they say something, and have one or two goals that you can 'work on.' I've already got my dogs and ponies all lined up," he had noted with a grin.

Even when leaders are committed to professional development and personal improvement, testing creates ambivalence. Will the test truly capture their strengths and weaknesses? Will the results make them look bad? Will it make any difference one way or the other? More than

any other factor, what happens *after* the test will determine whether it becomes a meaningful contribution to their development or an empty exercise that is quickly forgotten. Chappelow, commenting on "drive-by assessments," has noted that "providing assessment without developmental planning and followup almost guarantees that the organization does not get its money's worth."

A worst-case scenario is when district officials report the assessment results to supervisors but do not bother to inform the participants themselves. District leaders may feel that some kind of accountability function has been satisfied, but the procedure will do little to improve the quality of leadership.

More commonly, participants learn the results but are left to draw their own conclusions and take appropriate action. With motivated leaders and well-designed followup materials, this approach sometimes leads to improvement, but equally often the process gets sidetracked as busy administrators put it aside "until I have the time."

The best-case scenario is a thoughtful, structured process that provides ample support to participants as they interpret the results and fine-tune their professional-development plans. While districts vary in the resources they can put behind an assessment program, this chapter is based on the assumption that a carefully structured and well-supported followup process will yield maximum benefits both for leaders and the districts they serve.

With a well-chosen instrument that matches local needs, a district is positioned to reap great benefit from the process of assessing leadership. However, administering the test is not the end of assessment; indeed, the real work has just begun. District officials face four key steps in following up on the assessment: selecting the facilitator, choosing the format or setting in which the feedback will be provided, delivering the feedback clearly and thoughtfully, and translating the results into workable plans for improvement.

Selecting the Facilitator

Choosing a good facilitator is critical. "That person sets the tone, serves as the primary resource to help people understand their feedback, and assists them in developing strategies for overcoming any obstacles to meeting their development targets" (Richard Lepsinger and Antoinette Lucia 1997). At a minimum, the person chosen should have at least three qualities.

1. *Knowledge of the instrument.* In some cases, vendors require that tests be administered and interpreted by people who have been specially trained and certified to use that particular instrument. The Principal Perceiver, for example, requires that facilitators receive both initial training and periodic updating to be certified. Even where this is not the case, participants have the right to expect that those who deliver the feedback are thoroughly familiar with the instrument and the meaning of the results.

2. *Sensitivity.* Being assessed is often an intimidating experience that can easily threaten confidence and self-esteem. Facilitators must be aware of these issues and, with sensitivity and insight, be able to help leaders confront discrepant information. They must also be ready to point out the positive aspects of the feedback being presented.

3. *Credibility.* Participants are likely to listen respectfully to facilitators who are knowledgeable and sensitive, but for the final critical step—applying the results—participants must feel that their guide understands their situation. Glibly advising a principal to maintain a "customer focus" with parents will merely lead to cynicism unless the facilitator recognizes the daily reality of parents who don't know (or care) how their child is doing in school. In other words, facilitators should understand the conditions under which participants are working.

While there may be a number of people within the district who meet these standards, an outside facilitator is sometimes the best choice. Internal facilitators—especially those with supervisory responsibility over the participants—may elicit cautious, face-saving discussion rather than honest, wide-open discussion. An empathetic but honest outsider is more likely to stimulate the right kind of reflection. (In fact, supervisors can become participants themselves, giving them a chance to model reflective thinking and also sending a powerful message about the importance of assessment.)

Deciding on the Setting

The next step is to determine how the feedback will be provided to participants. Lepsinger and Lucia describe three possible formats.

1. *Group sessions.* If a district is involving many of its administrators in leadership assessment, the results can be presented in a group workshop. A group setting is not only cost-effective, it also can strengthen the feedback process by assuring leaders that they are not

alone and that their peers (even those they respect highly) may also have areas of weakness. In addition, group discussion can stimulate ideas that individuals acting alone might overlook.

One potential disadvantage to group sessions is the discomfort leaders may feel in talking publicly about perceived weaknesses. Depending on the organizational climate, they may be more concerned with keeping face than with engaging in an honest appraisal of their strengths and weaknesses. In addition, a group setting limits the individual attention any one leader can receive; issues that are important to some participants may never be discussed.

2. *One-on-one-settings.* Another common approach is to meet with each leader individually, allowing full attention to his or her needs. One-on-one sessions also create an atmosphere of confidentiality, making it easier to talk about weaknesses and the need for change.

However, person-to-person meetings may be more intense, since the participant receives the full, unwavering attention of the facilitator. Lepsinger and Lucia note that "the facilitator must have a sure sense of when to confront a recipient with painful truths and when to back off."

One-on-one sessions are also time-consuming for the facilitator, who may easily spend two or three hours preparing and carrying out each meeting. When there are many participants, this approach can quickly become overwhelming, especially if the facilitator has other duties in the system. Compromises are possible, however; a facilitator could begin by meeting with all participants to present the general explanation and interpretation of the results, and then follow up with one-on-one sessions to consider individual issues.

3. *Self-study.* If neither group nor one-on-one sessions are available (perhaps due to lack of time), leaders can simply be given the results along with some followup material. (Many vendors provide self-guided booklets that provide interpretations of the results as well as reflective and goal-setting activities.) Self-study provides maximum flexibility in scheduling (it can even be done in the comfort of home, away from the usual work distractions). It may also allow leaders to feel autonomous and in control, avoiding the "going to the woodshed" flavor of guided sessions.

Nonetheless, self-study has notable disadvantages. On their own, busy leaders may tend to give the activity a low priority, delaying until the feedback has lost much of its value. Misinterpretations are

always possible, and leaders reflecting in isolation may lack the perspective that comes from interacting with others.

These decisions about who should present the feedback and in what setting should not be allowed to obscure another key issue: timeliness. Districts conducting leadership assessments should ensure that feedback sessions are conducted as soon as possible after the instrument has been administered. As with any test, quick feedback is important; initially high interest tends to diminish as weeks or even months pass without any word of the results. Participants may even have trouble remembering the issues that were raised as they took the test.

Finally, feedback should be delivered at a time and in a place where participants can focus their entire attention on the results. Group sessions can be scheduled for a morning or afternoon away from the work sites with their paperwork, phone calls, and other distractions.

Delivering the Feedback

The moment of truth in any leadership-assessment program is the delivery of the data obtained by the instrument. Will participants be able to understand the data's meaning and grasp their implications? Test data can be complex; emotions often run high; and the implications are not always immediately clear. Van Velsor has suggested a simple rubric to sort through the key issues by asking three kinds of questions: What? So what? and Now what? This section addresses the first two questions, and the following one suggests a process for converting the data into a workable plan for professional development.

What Do the Data Say?

The first task in feedback sessions is making sure that participants understand what the results are saying. Although many assessment instruments provide a variety of interpretive aids, the amount of information can be overwhelming. In particular, participants may need assistance in understanding the nature of the item scales, the meaning of the feedback form, the norms being used, and the breakdown of individual items.

Understanding item scales. Typically, the items on a leadership instrument take one of two forms. Some are *mastery* items, asking respondents to indicate how well a certain skill is performed. For example, the scale may run from 1 (little skill) to 5 (great skill) on items such as:

- communicates expectations clearly
- conveys respect for teachers' ideas
- is well organized

Others are *frequency* scales, asking respondents to indicate how often a behavior is observed. For example, on the items mentioned above, the scale may run from 1 (seldom) to 5 (very frequently).

Van Velsor notes that when using 360-degree feedback, frequency scales have the advantage of focusing on behavior that has been observed, making the assessment somewhat more objective (and thus more easily accepted). However, she also points out that frequency scales have the disadvantage of suggesting that more is better, which is not always true. "Paying attention to detail" is generally a positive behavior, but at some point it turns into micromanagement.

Another issue is the number of possible choices. Odd-numbered scales (for example, 1 to 5) typically leave room for neutral or indecisive responses in the middle, leaving an escape hatch for raters who don't want to put much thought into the process (Van Velsor). Even-numbered scales tend to nudge respondents into committing to one position or the other.

A final question is whether the scale permits a response such as "not applicable" or "no opportunity to observe." Lacking such options, respondents may be forced to reply without having a firm basis for their rating. On the other hand, they may find it easier to avoid taking positions on questions they *are* in a position to answer.

Test users will find few instruments that strike a perfect balance among these questions, but the issues should be raised during the feedback process to help users better understand the meaning of their scores.

Interpreting feedback forms. Leadership instruments vary greatly in how they report results to test takers. Typically, there is some combination of numerical scores, narrative explanations, and visual representations such as charts, graphs, and grids. While these reports are often designed to be self-explanatory, participants can still benefit from clear and concise demonstrations of how their data are mapped out.

Understanding norms. In many cases, scores are stated in terms of comparison to some reference group. This is helpful information, since leaders (like everyone else) gauge their effectiveness by comparing themselves to others. We often don't know how good we are until we know how good others are. At the same time, it is important to keep in mind with whom we're being compared. If rank beginners are being compared with veterans, the gap could be discouraging unless they understand that their lower standing reflects differences in experience.

This issue is especially important for school leaders, since many leadership instruments are standardized on business executives. While there are undoubtedly many similarities between leadership in business and leadership in schools, there are also powerful differences. For example, it makes sense to measure business leaders on their "customer orientation." For school leaders, keeping customers happy may require a very different set of skills, even after they figure out who the customers are.

In many cases, the norm group is everybody who has taken the test—often a sizable group. While this gives an instant comparison with a large number of peers, this group may not be representative of the leadership population at large. Most likely, the ones who have taken the instruments are leaders who are especially interested in self-evaluation and professional development.

It's also important to note that, on many instruments, comparison to the norm is just a comparison. It lets leaders know where they fit in with other leaders, but unless there are studies showing that success on the test is correlated to success on the job, one has to be careful in drawing conclusions.

An alternative used on a few instruments is comparison to an *ideal score*. That is, test makers have established a hypothetical "perfect score" and have compared the leader's results to that ideal. Van Velsor notes that comparison to an ideal provides a sense of challenge for ambitious leaders, but others may be discouraged by their inevitably less-than-perfect results.

Analyzing the item breakdown. Test feedback typically presents scores on a number of general dimensions such as "communicates expectations clearly," "judgment," or "decisiveness." These categories are useful in identifying underlying themes in one's approach to leadership, but they can seem rather abstract. Analysis of *individual* items gives a more tangible picture of the specific behaviors that led to the overall score. For example, it is worthwhile knowing that one

has an average score in "vision," but even more valuable to see that "communicates the vision" is high while "encourages others to take ownership of the vision" is low.

Analyzing rater breakdown. On tests that involve 360-degree feedback, separate scores are given for the different categories of rater, such as superior, peer, and subordinate. Discrepancies among these groups can spark considerable interest and reflection, and the scores should be given careful attention. Van Velsor notes that while participants sometimes give greater weight to the opinions of supervisors, perceptions of all raters are potentially valuable, since each sees different facets of the leader's behavior. These contrasts are especially informative when broken out as the concrete behaviors typically described by individual items.

What Do the Data Mean?

Once the scores are understood, attention turns to their implications. While most instruments are based on theory or research that implies certain leadership behaviors are more effective than others, assessments will almost never end up with a simplistic judgment of "good leader" or "poor leader." Instead, the results will show a combination of strengths and weaknesses, some of which are critical issues for the leaders to address, while others have only a marginal impact on success.

For that reason, participants need to sort through the available data with some care, neither rejecting the findings out of hand nor embracing them uncritically. Data are simply data; they require human judgment to determine their significance.

To illustrate, we will imagine that Anna Sebring, the principal described at the beginning of the chapter, has received 360-degree feedback on the behavior "promotes organizational learning." Table 3 shows the ten items on this particular scale; table 4 shows her scores.

What conclusions should Sebring draw from these results? Lepsinger and Lucia suggest that the first step is to determine which results reflect strengths and which point to weaknesses. On instruments that are normed, a strength might reasonably be defined as any score in the upper quartile and a weakness as any score in the lower quartile. This test, however, is not normed; rather, it uses an "ideal score" method, assuming that 5 is the goal. Users are left to themselves to decide the lines between strong, average, and weak.

TABLE 3

Items That Measure the Behavior 'Promotes Organizational Learning'

1. The principal has a distinct vision for the future of this school.
   1———2———3———4———5
2. The principal frequently talks about ways we can reach our vision.
   1———2———3———4———5
3. The principal extends ownership of the vision to everyone.
   1———2———3———4———5
4. The principal recognizes and celebrates innovative efforts.
   1———2———3———4———5
5. The principal provides funding for professional-development activities.
   1———2———3———4———5
6. The principal encourages us to visit other classrooms.
   1———2———3———4———5
7. The principal gives teachers considerable leeway to make instructional decisions.
   1———2———3———4———5
8. When teachers in this school want to try a new method, they don't feel they have to get permission from the office.
   1———2———3———4———5
9. The principal encourages discussion of new ideas at faculty meetings.
   1———2———3———4———5
10. The principal is always putting articles and clippings about new ideas in teachers' mailboxes.
    1———2———3———4———5

TABLE 4

Results of 360-Degree Feedback on the Behavior 'Promotes Organizational Learning'

Items	*Self-Rating*	*Teachers' Rating*	*Peers' Rating*	*Supervisors' Rating*
1. Has a distinct vision for the future of this school	5.0	2.3	4.5	5.0
2. Always talks about ways to achieve the vision	3.0	2.1	4.4	5.0
3. Extends ownership of the vision to everyone	5.0	2.5	3.5	4.0
4. Recognizes and celebrates innovative efforts	4.0	4.5	4.6	5.0
5. Provides funding for professional-development activities	4.0	2.1	3.7	5.0
6. Encourages teachers to visit other classrooms	3.0	1.5	2.8	4.0
7. Gives teachers leeway to make instructional decisions	4.0	4.6	3.8	4.0
8. Encourages discussion of new ideas at faculty meetings	5.0	3.1	4.4	5.0
9. Frequently puts articles about new ideas in teacher mailboxes	3.0	2.1	3.2	4.0
10. Makes teachers feel their contributions are appreciated	4.0	4.5	3.5	4.0

Obviously, different participants may apply different standards to their results. Like many educators, Sebring automatically equates any five-point scale with the traditional grading system, so that a 5 becomes an A, a 4 is a B, and so forth. So for her, anything over a 4 (B) is a definite strength, while anything under 2.5 (C-) is cause for concern. At least initially, facilitators should accept participants' own classifications, since the ultimate goal is to understand, not to label. Unrealistic perceptions can be gently confronted as the discussion proceeds.

As Sebring looks at the results, she is relieved that her peers and supervisors rate her performance rather high, with no scores that could be considered weak. In thinking about this, she realizes that they have not really had a chance to see her in action. Because the assessment instrument did not have a "not observed" response, her colleagues could only judge her by their perceptions of her conversations and meetings with them. Thus, their ratings may have been an indirect reflection of her own perceptions.

In the teachers' ratings, four items fall into Sebring's self-defined "weak" category:

- always talks about ways to achieve the vision
- provides funding for professional-development activities
- encourages teachers to visit other classrooms
- frequently puts articles about new ideas in teachers' mailboxes

Strength areas identified by teachers include:

- recognizes and celebrates innovative efforts
- gives teachers leeway to make instructional decisions
- makes teachers feel their contributions are valued

Lepsinger and Lucia suggest that, in considering weaknesses and strengths, participants distinguish between behaviors that will have a high impact on the organization and those that will have less impact. Sebring's strengths and weaknesses, as well as several areas that she is not sure about, are displayed in table 5.

Sebring can find at least one weakness she considers low-impact: not putting articles in teachers' mailboxes. Recent staff reductions combined with a new high-stakes state assessment program have kept her teachers near exhaustion; they barely have time to read unsolicited articles, much less act on them. In any case, her veteran staff members, having seen heralded reforms come and go, are inclined to

TABLE 5

Separation of Strengths and Weaknesses into High- and Low-Impact Categories

	More important	*Less important*
Strengths	Recognizes and celebrates innovative efforts	
	Gives teachers leeway to make instructional decisions	
	Makes teachers feel their contributions are appreciated	
Weaknesses	Has a distinct vision for the future of this school	Encourages teachers to visit other classrooms
	Always talks about ways to achieve the vision	Frequently puts articles about new ideas in teacher mailboxes
	Extends ownership of the vision to everyone	
Unclear areas	Why is faculty unclear about my vision?	
	Why does faculty not recognize my efforts in funding professional development?	
	Why don't teachers see faculty meetings as a forum for new ideas?	

skepticism about the latest bandwagons. They do seem receptive to concrete local examples of new ideas, and Sebring realizes she has shared a number of such ideas in hallway conversations (a tactic that this assessment doesn't recognize). All things considered, she does not see this particular "weakness" as a major problem.

In addition, the low score on encouraging teachers to visit other classrooms does not strike her as a high-impact issue. It's a nice idea but probably not the kind of thing that can transform a school.

This initial classification of strengths and weaknesses should be followed by one other step: articulating issues that need to be clarified (Lepsinger and Lucia). The meaning of the feedback is not always clear, and participants should be encouraged to identify any areas that are ambiguous or puzzling. Sebring indeed has several questions about her results:

- Why do teachers seem uncertain about her vision?
- Why do teachers rate her so poorly on her ability to find funding for professional-development activities?
- What explains the difference between her perceptions and the faculty's perceptions about what happens at faculty meetings?

In considering the first question, she realizes that as a first-year principal she has tried to avoid coming on too strong until she gets a feel for the school. This initial hesitation has been reinforced by the fact that her predecessor had been an energetic promoter of "vision" and had enticed the staff into spending many hours developing their "twenty-first century dreams," only to blithely ignore the resulting product. Anna had realized early in the year that the last thing these cynical teachers wanted to hear was more rhetoric about vision. Thus, while she had a distinctive vision for the school, and shared it freely with her administrative colleagues, she had been soft-pedaling her ideas when talking with her teachers.

She was hurt by the teachers' low rating on the funding question, since she had spent many hours lobbying, negotiating, and just plain begging for her school's fair share of the money. After thinking about it, however, she realized that most of this effort had gone on behind the scenes. Her sense of professionalism makes her reluctant to talk about district politics with her faculty, so they can judge her only by the results, which are admittedly meager; the district is simply strapped for cash this year.

The question of faculty meetings is more puzzling. She believes that meetings should be used for professional dialogue rather than

routine announcements. In her view, the monthly meetings had been rather stimulating. After discussing this with the facilitator, she decides that she will have to talk quietly with a few teachers to find out why their perceptions are so different from hers.

After dealing with these issues, Sebring and the facilitator discuss the implications of the assessment. Three issues seem to stand out. First, she has a real strength in making teachers feel valued, accepted, and respected as autonomous professionals. In fact, their ratings in this area were consistently higher than her self-assessment. She is pleased by this finding, since affirmation and support are important values for her.

Second, it appears that she has not conveyed her vision to the staff. While she still defends her decision not to freely dispense visionary rhetoric, she also recognizes the importance of collectively facing questions about the school's future. She knows that the state assessment is a new way of doing business, and simply reacting to mandates will not get the school where it needs to go. She also realizes that the high scores for respecting teacher autonomy may not be an unalloyed strength: at some point, "respecting autonomy" may become "laissez-faire."

Finally, she reconsidered her lowest score ("encouraging teachers to visit other classrooms"). Although she had initially ranked this as a low-impact item, she now wondered if it represented a deeper issue. Early in the year she had suggested such visits to a few teachers, but backed off when they were immediately dismissive. Although she still didn't think that mandating visits would accomplish much, she now wondered if she had dropped the idea too quickly. Thinking about it, she realized that the faculty had often displayed a "live and let live" philosophy that frowned on the notion that some teachers might be better than other teachers. Again she was faced with the realization that her desire for a harmonious school might be eliminating the kind of creative conflict that leads to learning.

This example should make it clear that interpreting feedback is a much more subtle process than simply getting a thumbs-up or thumbs-down on a checklist of traits. What initially appears to be a strength may contain elements of weakness, and vice-versa. Some issues may remain ambiguous, and even obvious strengths or weaknesses may be relatively unimportant in the larger scheme of things. In evaluating assessment results, both facilitators and participants should recognize

that they are not so much arriving at a conclusion as setting out on an open-ended journey.

Forming a Development Plan

All this activity is little more than an intellectual exercise unless leaders are able to convert their results into a plan for improvement. If there are weaknesses, those need to be addressed. If there are strengths, those should be exploited. If there are uncertainties, those should be explored further.

For that reason, participants in assessment should be encouraged to develop an action plan as soon as possible. Lepsinger and Lucia say, "If participants do not take meaningful steps to translate their feedback into action within two weeks of leaving the work session, they will probably never do so." Doing so involves six key steps.

1. *"Consolidating" the feedback.* Lepsinger and Lucia note that leadership assessments are typically rich in data and insights. It may take several days just to assimilate what has been learned, without even getting started on the action plan. Facilitators should not plan to leap from feedback to action in one meeting.

2. *Setting development goals.* At this point the feedback analysis, as insightful as it may be, needs to be shaped into an agenda for action. The plan should be couched in specific behavioral terms; that is, What must the leader *do* to build leadership capacity? (Banal generalities such as "be more decisive" and "improve communication" are rarely helpful.) To continue the example begun earlier, Anna Sebring decides that her top priority is leading her staff to *collectively* address the challenges facing them. That is, she must find ways to engage teachers in dialogue about future directions and get them to collaborate on curricular and instructional decisions, even if this occasionally sparks conflict.

Chappelow suggests three questions that may be helpful in choosing a goal: "Does the goal motivate and energize me? Will achieving this goal help me be more effective in my current position? Will my organization benefit from this goal?"

3. *Choosing a strategy.* Development goals aim to take participants to places they have never been before, so a search for creative strategies is essential. Lepsinger and Lucia suggest five possible approaches.

First, participants may find what they need in the rich professional literature on leadership and change. Although much of this material is driven by fads and self-promotion, it contains some thoughtful advice and genuine nuggets of wisdom. Test publishers often include bibliographies tailored to the issues raised by their instruments, and several professional journals contain helpful reviews of relevant books. A newer approach is to join an administrative LISTSERV such as those operated by the American Association of School Administrators and the ERIC Clearinghouse on Educational Management.* These discussion groups make it easier to find other administrators facing the same problem.

Second, participants may engage in self-monitoring. After identifying a key behavior to strengthen, they simply chart the number of times they use it. Over time, this tracking process keeps leaders on target and makes the skill seem increasingly natural.

Third, participants can work with a coach, mentor, or consultant. When a desired behavior is complex, self-monitoring is less effective than being exposed to modeling and dialogue with an experienced, knowledgeable guide.

Fourth, specific training programs and seminars are plentiful, especially those dealing with well-defined skills such as delegating, time management, and team-building.

Finally, on-the-job assignments can enhance targeted skills. For example, a principal interested in developing his team-building skills could take on responsibility for a districtwide curriculum committee.

4. *Writing a development plan.* Lepsinger and Lucia emphasize the importance of putting the development plan into tangible form. An effective written plan will include a clear statement of the goal, the standards for measuring success, the change strategies, the related action steps, and key resource people.

* Linked by the K12ADMIN listserv, administrators can share ideas, help each other solve problems, exchange information, and make contacts with geographically remote colleagues. In addition to facilitating dialogue among subscribers, K12ADMIN provides an avenue for sharing information about such things as upcoming conferences and useful publications.

There is no charge to subscribe to K12ADMIN. To subscribe, send the message **subscribe K12ADMIN Your Name** (e.g., K12ADMIN Jane Doe) to the following email address: **listserv@listserv.syr.edu**

For additional information about K12ADMIN, contact Mary Lou Finne, user services coordinator, ERIC Clearinghouse on Educational Management, University of Oregon. Voice: (541) 346-2333. Email: mlfinne@eric.uoregon.edu

5. *Monitoring the plan.* Participants should be encouraged (or required) to take stock periodically, reviewing their progress, identifying difficulties, and modifying strategies where appropriate.

6. *Supporting the development process.* This kind of professional development is not inherently expensive, but districts nonetheless should be ready to provide any resources that may be necessary. Having invested money in the assessment and the delivery of feedback, it makes little sense to tie the purse strings. Participants who are left to fend for themselves will hear the message (intended or not) that assessment is a ritual rather than a chance for real change.

Searching for Ground Truth

Getting direct, unvarnished feedback about leadership capabilities is a nervous experience for almost everyone, even highly capable leaders. In fact, Lepsinger and Lucia note that the strongest leaders—those with unquestioned belief in themselves and their skills—may find it most difficult to have those beliefs challenged.

School leaders, who are always under pressure to accomplish a great deal with limited resources, are especially vulnerable. Confronting a wide array of ambiguous issues, and facing multiple constituencies with conflicting agendas, they often place a premium on appearing strong and in control.

Kaplan and Paulus note that executives—like everyone else—put a lot of energy into maintaining their self-image. They may be willing to make fine adjustments, but major changes are likely to be resisted. For example, a principal who prides himself on listening to others might be willing to admit that on a particular occasion he failed to listen to a teacher, but would probably resist evidence that a majority of his staff members considers him to be a poor listener. Or, he may rationalize away the evidence by saying, "They don't understand what I'm trying to do."

Dealing with such concerns requires a thoughtful, sensitive facilitator who can provide participants with positive affirmation while also challenging them to improve. Several guidelines may be helpful.

1. *Listen first.* Lepsinger and Lucia argue that directly challenging participants' perceptions often leads them to dig in their heels. Preferably, facilitators should listen to arguments and rationalizations without immediately directly challenging them. "By letting recipients vent, you are clearing the way for them to acknowledge

and accept at least a few of the messages they received and thereby increasing the likelihood that they will act on them."

This advice is reinforced by the simple fact that score interpretations *are* sometimes ambiguous, and that apparent weaknesses sometimes represent a best-possible response to the leader's unique circumstances. For instance, a test of visionary leadership might consider it a weakness that a principal fails to articulate and frequently discuss the school's vision for the future. Yet if the previous principal talked *ad nauseum* about vision but never followed through, the staff may be understandably cynical about this kind of happy talk, and the new principal may be wise to tone down the rhetoric and build commitment to vision in other ways.

2. *Focus on what can be learned rather than on the raw scores.* Test scores are merely reflections of underlying behavior, not ends in themselves. Knowing how one compares with other leaders is not as valuable as knowing how one can become more effective on the job. Participants should leave the session with heightened understanding of their own approach to leadership as well as a plan for continuing development.

3. *Focus on strengths as well as weaknesses.* People often undergo assessment using a deficit model that assumes the goal is to fix what's broken. In reality, improvement can come from leveraging strengths as well as remedying weaknesses. For example, a leader whose assessment shows strong interpersonal skills may be encouraged to begin consciously using those abilities in new ways. In addition, zeroing in on strengths can alleviate some of the anxiety normally found in receiving feedback, and this positive approach can help convince the participant that the test is credible.

4. *Focus on behavior rather than abstract qualities.* Although people are often quick to reduce leadership to convenient short-hand abstractions ("communication," "vision," "decisiveness"), those broad categories by themselves are not very informative. If an assessment indicates that you are below average on vision, how do you become above average? What specific attitudes and behaviors have to change?

In their book *Hope Is Not a Method*, Gordon Sullivan and Michael Harper describe the journey the U.S. Army has taken to become a learning organization. At the heart of the process is the willingness to face the brutal facts—what the authors call "ground truth." Denying *what is* will not lead to *what should be*. Only pervasive, persistent evaluation will let us know when we go off course.

While it is always painful to confront our own failures and inadequacies, the payoff for this kind of honesty is huge: an improved capacity to do the job we were meant to do. Becoming better equipped to help children achieve their goals is worth a few temporary dents in our self-esteem.

FORETASTE

No matter how you define leadership, you can probably find an instrument to match. The assessments surveyed in this chapter represent a modest sampling of the instruments available.

Benchmarks

COMPASS: The Managerial Practices Survey

The Comprehensive Leader

Educational Administrative Style Diagnosis Test (EASDT)

Educational Administrator Effectiveness Profile

Leader Behavior Analysis II (LBAII)

Leadership Competency Inventory (LCI)

Leadership/Impact™

Leadership Orientations

Leadership Practices Inventory (LPI)

Leadership Skills

Leadership Sphere Inventory (LSI)

Leadership Workstyles

Life Styles Inventory (LSI)

Myers-Briggs Type Indicator (MBTI)

NASSP Assessment Centers

The Principal Perceiver

Professional Development Inventory

Profile of Aptitude for Leadership (PAL)

The Survey of Leadership Practices

The Survey of Management Practices

Instrument Profiles

The instruments described in this chapter represent the range and varieties of leadership assessment available to school districts, with emphasis on products from commercial vendors, since these are most likely to come with support materials.

These capsule summaries are not intended to be rigorous analyses of either technical quality or suitability, nor should their inclusion be construed as an endorsement of their use. As frequently noted in the preceding chapters, potential users are the only ones who can decide which test best serves the needs of their schools. Instead, this listing is designed to provide school leaders with a convenient survey to begin their search.

Readers seeking a more indepth review will find some of these instruments described in greater detail in *Feedback for Managers: A Review and Comparison of Multi-rater Instruments for Management Development* (Jean Brittain Leslie and John W. Fleenor).

Benchmarks

Center for Creative Leadership
P.O. Box 26300
Greensboro, NC 27438-6300
Phone: 336-288-7210
Fax: 336-288-3999
www.ccl.org

Purpose

Benchmarks is a professional-development instrument designed to measure strengths and weaknesses of executives.

Target Group

Middle- and upper-level executives.

Description

Benchmarks consists of 164 items. The largest section yields feedback in sixteen categories. Four are in the area of "meeting job challenges": resourcefulness, doing whatever it takes, being a quick study, and decisiveness.

Five are in the area of "leading people": leading employees, setting a developmental climate, confronting problem employees, work team orientation, and hiring talented staff.

Seven are in the area of "respecting self and others": building and mending relationships, compassion and sensitivity, straightforwardness and composure, balance between personal life and work, self-awareness, putting people at ease, and acting with flexibility.

Another section generates feedback on six "problems that can stall a career": problems with interpersonal relationships, difficulty in molding a staff, difficulty in making strategic transitions, lack of follow-through, overdependent, and strategic differences with management.

A third section yields information on how the leader handles a variety of typical business assignments, and the final section asks raters to identify which eight (out of sixteen) success factors are most important in their organization.

Response forms are provided for the leader and eleven observers (supervisors, peers, and subordinates). Responses are on a five-point Likert scale, ranging from "not at all" to "a very great extent" on the first section. The scale on the second section (derailment factors) ranges from "strongly disagree" to "strongly agree," while the third section (ability to handle various jobs) ranges from "among the worst" to "among the best."

Feedback

Participants are given a feedback report divided into three sections: leadership skills and perspectives, problems that can stall a career, and handling challenging jobs. Feedback includes average ratings by self and others; comparison of ratings to norm groups; importance ratings; and an item-by-item breakdown. The report also highlights the fifteen highest and lowest rated items within each rating group, as well as items where the range of responses was three points or higher within the same group of raters. Norms are based on

a sample of high-level or mid-level managers, depending on the level of the person being rated.

Followup

The publishers provide a developmental learning guide that helps participants analyze their results, set a developmental goal, choose a strategy, and implement the plan. The Center for Creative Leadership also offers developmental workshops.

Theory and Rationale

Benchmarks is based on research on the developmental experiences of business managers in Fortune-500 firms. Interviews and surveys asked executives to describe key events in their managerial careers and what they had learned from these experiences. The results led to sixteen categories of key developmental events, as well as six factors that can cause "derailment."

Administration

Tests can be self-administered. Scoring is done by the publisher, though scoring software can be licensed. The Center for Creative Leadership requires certification for trainers and facilitators who will be giving feedback.

Statistical Validation

Not provided with review materials.

Uses

Most appropriately used for professional development of experienced managers.

Cost

A set of Benchmarks instruments, which includes twelve surveys to assess one person, costs $245.

COMPASS: The Managerial Practices Survey

Manus Associates
100 Prospect Street
South Tower
Stamford, CT 06901
Phone: 800-445-0942
Fax: 203-326-3890
Email: manus1@aol.com

www.rightmanus.com
Contact: Debbie Horne

Purpose

COMPASS: The Managerial Practices Survey provides information on current leadership behavior to assist in professional development.

Target Group

Managers at all levels in business, public, and military organizations.

Description

COMPASS consists of seventy items providing feedback in four clusters containing fourteen scales: communicating (informing, clarifying, monitoring), decision-making (planning, problem-solving, consulting, delegating), motivating (inspiring, recognizing, rewarding), and building relationships (supporting, mentoring, networking, teambuilding). Respondents are asked to rate behavior, ranging from 1 ("never, not at all") to 4 ("usually, to a great extent"), with "not applicable" and "don't know" responses allowed.

Leaders rate themselves and are also provided feedback from up to eight peers and subordinates. Leaders and their supervisors are also asked to rate the importance of each category for the particular setting.

Feedback

Results are reported numerically and graphically. Feedback includes an overall score for each scale as well as results for individual items, with side-by-side comparisons of assessments by self, colleagues, and subordinates ("direct reports"). Feedback on delegating, rewarding, and mentoring is provided only from subordinates. Importance ratings by supervisor and self show the relative importance of each category for the work environment.

Followup

Development and planning guides are available to help participants understand the fourteen practices and put them into action. Developmental workshops are also available from the publisher.

Theory and Rationale

COMPASS was developed from an extensive research program headed by Gary Yukl, using factor analysis, judgment, and deduction

to create a taxonomy of behaviors related to managerial effectiveness.

Administration

COMPASS is self-administered, taking about twenty to thirty minutes to complete. Scoring is done by the publisher and reported in about two weeks. Users must be certified by Manus, which offers certification workshops.

Statistical Validation

Ratings of internal consistency and test-retest reliability are high, with internal reliability ranging from moderate to high. Studies indicate that some of the scales correlate significantly with performance on the job; three scales (clarifying, monitoring, and networking) have been found to correlate with effectiveness for elementary principals.

Uses

COMPASS is most appropriately used as a professional-development tool to help leaders understand and improve their managerial practices.

Cost

COMPASS is priced at $225 a set for 1-50 sets, $200 for 51-100 sets, $180 for more than 100 sets. A set consists of a self-assessment instrument, eight copies of peer and subordinate instruments, an importance questionnaire for the supervisor, publisher scoring of the results, a computer-generated feedback report, and a manual.

The Comprehensive Leader

HRDQ
2002 Renaissance Boulevard #100
King of Prussia, PA 19406-2756
Phone: 800-633-4533
Fax: 800-633-3683
www.hrdq.com
Contact: Laurie Ribble Libove (LRLIBOVE@HRDQ.com)

Purpose

The Comprehensive Leader is designed to help participants identify their strengths in the area of strategic and visionary leadership and formulate professional-development plans.

Target Group

Leaders in many types of organizations, profit or nonprofit. The assessment is not limited to those with formal supervisory authority over the people they are leading.

Description

The Comprehensive Leader consists of forty items assessing the leader in four areas: knowledge of self, knowledge of others, knowledge of the organization, and knowledge of the world. Respondents are asked to indicate the degree to which the statement is true of them: completely characteristic, mostly characteristic, somewhat characteristic, mostly uncharacteristic, or completely uncharacteristic. Related surveys are available for observers (peers, subordinates, or supervisors).

Feedback

Test-takers are given scores in each of the four dimensions, ranging from 10-50; scores of 40 or above are considered to be "relative strengths." (If no dimension has a score above 40, the one with the highest subtotal is designated a relative strength.) The publishers note that norms have not yet been established and that the cutoff score of 40 should be regarded as an estimate.

Scores may show strengths in from one to four dimensions, in differing combinations. For example, a participant may have strengths in "knowledge of self" and "knowledge of the organization," or in "knowledge of organizations" and "knowledge of the world." For each of the fifteen possibilities there is a profile consisting of a one-paragraph description and several questions focused on continuing growth.

Followup

Participants receive a booklet containing explanations of the scores and suggestions for acting on the insights gained through the assessment. A facilitator's guide includes instructions for presenting a feedback session, transparency masters, and additional development activities.

Theory and Rationale

The Comprehensive Leader was designed as an "easy-to-train" model focused on the essentials of visionary leadership. A review of the leadership literature identified more than 150 key leadership behaviors that were sorted into categories. After additional analysis, the test developers formulated a leadership model around four major

dimensions: knowledge of self, knowledge of others, knowledge of the organization, and knowledge of the world. The assumption is that the leader's awareness of these four dimensions is the root of personal conviction and earned credibility.

Administration

The instrument can be self-administered and self-scored. The instrument takes about ten minutes, scoring five to ten minutes, profile development about fifteen minutes.

Statistical Validation

Not provided with review materials. The publishers note that data collection is ongoing.

Uses

Most appropriately used for leadership development.

Cost

A preview pack containing test booklet, feedback form, and facilitator's handbook is available for $45. Five-packs of both the participant form and feedback form cost $40, with quantity pricing available.

Educational Administrative Style Diagnosis Test (EASDT)

W. J. Reddin and Associates
P.O. Box 324
Fredericton, New Brunswick
Canada E3B 4Y9
Phone: 506-452-7194
Fax: 506-452-2931

Purpose

The Educational Administrative Style Diagnosis Test is designed to identify and stimulate thinking about the styles of educational administrators.

Target Group

The EASDT is designed solely for use by educational administrators.

Description

The instrument consists of fifty-six forced-choice items probing various dimensions of administrative style. The instrument assesses task-orientation, relationship-orientation, and effectiveness (the appropriateness of a particular style in a given situation). The interaction among these three dimensions leads to eight possible styles: missionary, compromiser, developer, executive, deserter, autocrat, bureaucrat, and benevolent autocrat.

Feedback

Results are self-scored and self-recorded. The results include scores for each of the eight styles; higher scores indicate that the participant leans toward that style, while very high scores indicate a dominant style. Results can be graphed for a quick overview, and participants are also provided with an overall "synthesis" score. Brief interpretations of the different styles are provided. An associated Test Interpretation Manual is available.

Followup

Not provided with review materials. A guide for facilitators suggests several formats for helping participants understand the results.

Theory and Rationale

The EASDT is based on the "3-D Theory of Administrative Effectiveness," which uses the widely recognized concepts of task-orientation and relationship-orientation from a situational point of view.

Administration

The EASDT can be self-administered and self-scored in about thirty minutes.

Statistical Validation

Not included with review materials. The EASDT has been frequently used in dissertation research.

Uses

The EASDT is designed solely for use in training and development, and should not be used as an appraisal instrument.

Cost

The publishers will make available photocopies for free use for educational purposes only. Otherwise, the instrument is no longer available.

Educational Administrator Effectiveness Profile

Human Synergistics
39819 Plymouth Road
Plymouth, MI 48170
Phone: 734-459-1030
Fax: 734-459-5557
www.humansyn.com

Purpose

This Educational Administrator Effectiveness Profile is designed to help school administrators understand their managerial and leadership behaviors and develop their skills in those areas.

Target Group

K-12 school administrators.

Description

The instrument consists of 120 items using a seven-point Likert scale ranging from "almost never" to "always." Questionnaires are provided for the leader and for others (supervisors, peers, and/or subordinates.) Most of the questions measure eleven specific management skills: setting goals and objectives, planning, making decisions and solving problems, managing business and fiscal affairs, assessing progress, delegating responsibilities, communicating, building and maintaining relationships, demonstrating professional commitment, improving instruction, and developing staff.

In addition, ten questions lead to "summary perceptions" about leadership, including "have a vision of what could be and help others work toward it," "convince staff that their effort makes a difference," and "provide a work environment where people care about each other."

Feedback

Participants receive graphic and numeric feedback on summary perceptions and the eleven managerial skills (for "self" and "other" ratings). Item-by-item breakdowns are also available.

Followup

The publishers provide a self-development guide containing detailed discussions and interpretations for each of the eleven management skills, as well as advice on formulating an action plan for personal improvement.

Theory and Rationale

Not provided with review materials. The self-development guide notes that the eleven categories of managerial skill were developed through extensive research.

Administration

Not provided with review materials.

Statistical Validation

Not provided with review materials.

Uses

Most appropriate for leadership development in K-12 settings.

Cost

The cost is $45 for a package of the profile instruments.

Leader Behavior Analysis II (LBAII)

Blanchard Training and Development
125 State Place
Escondido, CA 92029
Phone: 800-728-6000
Fax: 619-489-8407
www.blanchardtraining.com

Purpose

The LBAII is designed to give leaders a better understanding of their leadership style by contrasting self-perceptions and other's perceptions (boss, associates, team members).

Target Group

Leaders and managers in all types of organizations.

Description

The LBAII consists of twenty hypothetical leadership situations to which participants are asked to choose the appropriate strategy from the four options listed. Forms are provided for the leader and others familiar with his or her work.

Feedback

The feedback provides numeric and graphic data showing how the leader and others rated the frequency with which each of the four LBAII styles is used. "Effectiveness scores" (based on the appropriateness of each response to the hypothetical situations) are also

computed. The feedback profile is organized around four questions: Do I see myself as others see me? Am I flexible? Do I manage people differently? Do I diagnose well?

Followup

The publishers offer a variety of books, articles, visual materials, and training programs centered on the concept of situational leadership. Training for trainers is also available.

Theory and Rationale

The LBAII is based on the Hersey and Blanchard theory of situational leadership, which assumes that the "right" leadership behavior depends on matching the appropriate style with the needs of employees.

Administration

The LBAII is administered onsite. It can be self-scored or can be computer-scored by the publisher with more feedback. (Onsite licenses are available for computer scoring.)

Statistical Validation

Not provided with review materials.

Uses

Most appropriate for professional development.

Cost

The LBAII costs $8.95 per package, including the assessment form and scoring materials.

Leadership Competency Inventory (LCI)

Hay/McBer
Training Resources Group
116 Huntington Avenue
Boston, MA 02116
Phone: 800-729-8074
Fax: 617-927-5060
Email: trg_mcber@haygroup.com
http://trgmcber.haygroup.com

Purpose

The LCI measures four key leadership competencies to assist in development of leadership capacity.

Target Group

The LCI is aimed at leaders in business and nonprofit organizations.

Description

The LCI consists of thirty-eight items that assess four leadership competencies: information seeking, conceptual thinking, strategic orientation, and customer-service orientation. Items assess behaviors using a five-point Likert scale, ranging from 0 ("never") to 4 ("extremely frequently"). Related questionnaires are provided for supervisors, peers, and/or subordinates.

Feedback

An interpretive booklet allows participants to profile and understand their scores. Ratings are expressed in terms of frequencies, with greater frequency seen as preferable for all four competencies. The booklet provides additional information on the *level* of competencies (that is, on how well they are performed).

Followup

The publisher provides a "leadership competency development assistant" that provides interpretations and extended examples of each competency, addresses obstacles to competency development, and provides resources for further development.

Theory and Rationale

Not provided with review materials.

Administration

Self-administered and scored.

Statistical Validation

Not provided with review materials.

Uses

Most appropriate for professional development.

Cost

$65 per package of 10 self-questionnaires, interpretive notes
$25 per package of 10 feedback (observer) questionnaires
$55 per package of 10 developmental guides

Leadership/Impact™

Human Synergistics
216 Campus Drive
Arlington Heights, IL 60004
Phone: 847-590-0995
Fax: 847-590-0997
Email: info@hscar.com
Internet: www.hscar.com

Purpose

Leadership/Impact is designed to show leaders how their personal styles affect their organizations.

Target Group

Leaders in all types of organization.

Description

Leadership/Impact is a 360-degree feedback instrument consisting of two main sections. "Impact on Others" consists of ninety-six questions assessing the degree to which the leader causes others to act in certain ways. Responses range from 0 ("not at all") to 4 ("to a very great extent"). The twelve categories are humanistic-encouraging, affiliative, approval, conventional, dependent, avoidance, oppositional, power, competitive, perfectionistic, achievement, and self-actualizing.

"Leadership Strategies" consists of sixty items assessing the degree to which the leader uses ten different strategies: envisioning, role modeling, mentoring, stimulating thinking, referring, monitoring, providing feedback, reinforcing, influencing, and creating a setting. Responses range from 0 ("never") to 4 ("always").

Forms are provided for self-assessment and assessment by others. In addition to these two sections, participants are asked to indicate the type of effect they *want* to have on others ("ideal impact").

Feedback

Participants receive verbal and graphic feedback in three main sections. First, they are presented with data showing their ideal impact compared with actual impact as reported by peers, associates, and supervisors.

Second, they receive feedback on "Impact on Others," which shows the degree to which others respond with constructive styles, passive/defensive styles, or aggressive/defensive styles. Scores are normed on a sample of 500 leaders. Item-by-item feedback is also provided, with direct comparisons of the leader's assessment and feedback from others.

Third, they receive data on "Leadership Strategies," showing the degree to which they use strategies that are prescriptive (encouraging others to take positive actions) or restrictive (restraining others from taking negative actions). Item-by-item feedback is provided for each of the ten strategies, with direct comparison of self-assessment and feedback from others.

Followup

Participants are provided with a detailed confidential feedback report, containing a section on "The Next Steps" with several goal-setting activities. Recommended readings are also listed.

Theory and Rationale

Leadership Impact is based on the assumption that leadership is best measured by the *impact* that leaders have on other people in their organizations.

Administration

Participants complete a paper-and-pencil survey form.

Statistical Validation

Not provided with review materials.

Uses

Leadership/Impact is most appropriate for professional development.

Cost

$175 per kit, which includes materials for self and eight others, scoring, and feedback report.

Leadership Orientations

Dr. Lee Bolman
Rm 308, Bloch
Bloch School of Business and Public Administration
University of Missouri

5110 Cherry Street
Kansas City, MO 64110-2201
Phone: 816-235-5407
Fax: 816-333-9775

Purpose

Leadership Orientations was developed by Lee Bolman and Terry Deal for use as a research instrument and a professional development aid.

Target Group

Leadership Orientations is suitable for leaders in a variety of domains.

Description

The first of three sections consists of thirty-two items on a five-point Likert-type scale; it asks participants to indicate how often each statement is true of them, from "never" to "always." The second section consists of six forced-choice questions asking participants to describe their style. The final section asks them to rate their effectiveness as leaders and managers in comparison with other individuals with similar levels of experience and responsibility. Similar versions are available for rating by others.

Feedback

No feedback materials were supplied with review materials. (The instrument was originally designed for research and is not being marketed as a commercial product.) However, participants can easily score their own results, which illumine their relative orientations on four main scales: structural, human resource, political, and symbolic. There are also subscales for analytic, supportive, powerful, inspirational, organized, participative, adroit, and charismatic.

Little explanatory material is provided with the instruments, but the authors have described the four main orientations in other publications; see, for example, Lee Bolman and Terry Deal (1991).

Followup

None available.

Theory and Rationale

Leadership Orientations is built around theory and research that has identified four "frames" that leaders can use to think about their work: *structural,* which emphasizes goals, planning, and coordination; *human resource,* which is sensitive to the human needs of

others: *political*, which recognizes the ways that people seek to advance their own interests; and *symbolic*, which focuses on the rituals, myths, and ceremonies that give meaning to organizational culture. The assumption is that all four frames play important roles in effective leadership.

Administration

Leadership Orientations can be self-administered and self-scored.

Statistical Validation

Extensive evidence of reliability is provided with the instrument.

Uses

Leadership Orientations has been used mainly for research, but it can be helpful in professional development if facilitated by someone familiar with the four-frames concept.

Cost

None. The authors routinely grant permission for noncommercial, research use of the instrument. They do ask that users provide them with copies of any research reports produced using results from the instrument.

Leadership Practices Inventory (LPI)

Jossey-Bass/Pfeifer
350 Sansome Street
5th Floor
San Francisco, CA 94104-1342
Phone: 800-274-4434
Fax: 800-569-0443
www.jbp.com

Purpose

"As you set out to train others, we ask that you share in our purpose—*to help liberate the leader in everyone*."

Target Group

Leaders at all levels in all organizations.

Description

The LPI consists of thirty items focused on five key leadership practices: challenging the process, inspiring a shared vision, enabling

others to act, modeling the way, and encouraging the heart. Items use a ten-point Likert scale ranging from 1 (almost never) to 10 (almost always). A self-report and an observer rating can be used by superiors, subordinates, peers, or others.

Feedback

The feedback provides overall ratings for the five dimensions of leadership, as well as a breakdown of individual items. Participants are presented with side-by-side comparisons of how their self-ratings compare with those of superiors, subordinates, and coworkers. The feedback includes percentile rankings using a norm group consisting of all leaders and observers who have taken the LPI since 1988. Participants are also given a rank-order listing of all items.

Theory and Rationale

The LPI is based on extensive research by James Kouzes and Barry Posner that focused on how "ordinary people accomplish extraordinary things." Leaders were asked to describe a "personal best" leadership experience. Their responses showed a consistent pattern that the researchers encapsulated in the five practices that are at the heart of the LPI.

Administration

Self-administered. The publishers advise that participants solicit the involvement of supervisors, subordinates, and peers who will be completing the observer form. Scoring can be done on-site by a facilitator, using the scoring software provided by the publishers.

Statistical Validation

The publishers report high reliability for the LPI, as well as excellent face validity. In addition, factor analysis studies show that the five practices are separate entities. A variety of studies have linked LPI performance with various measures of on-the-job success (examples are given in the facilitator's handbook).

Followup

The publishers provide a facilitator's guide that outlines a followup meeting to explain the results and begin the development process. Participants are given a handbook that interprets the scores and recommends activities for professional development. Under the headings "learning by doing," "learning from others," and "learning in the classroom or on your own," the workbook offers professional-development suggestions for each of the five practices.

Uses

The instrument is intended and most appropriate for professional-development activities.

Cost

The Leadership Practices Inventory—Individual Contributor (LPI-IC) *Facilitators Guide* includes one copy each of the self and observer instruments and scoring software. Quantity discounts are available.

LPI-IC Facilitators Guide—$49.95
LPI-IC: Self—$12.95
LPI-IC: Observer—$3.95

Leadership Skills

Acumen
4000 Civic Center Drive
5th Floor
San Rafael, CA 94903
Phone: 800-544-8626
Fax: 415-479-5358
Email: acumen@acumen.com
www.acumen.com
Contact: Jeannie Elrod

Purpose

Leadership Skills seeks to stimulate leadership development by informing leaders of their skills in major leadership domains.

Target Group

Leaders of organizations in a wide range of industries, including education.

Description

Leadership Skills is a multirater instrument that provides feedback on sixteen leadership skills in four domains: task management (informing, efficiency, planning, and problem-solving); team development (performance feedback, relationship skills, staff development, and team motivation); business values (quality improvement, customer focus, and promoting innovation); and leadership (accountability, empowerment, influence, mission skills, and networking).

The instrument contains 96 items using a 5-point Likert scale ranging from 1 ("never") to 5 ("always").

Feedback

Participants receive graphic and narrative feedback, including summary results and item-by-item breakouts comparing self-scores and ratings by others. Detailed interpretive information is provided.

Followup

The feedback report includes many activities and ideas for using the results for professional development.

Theory and Rationale

Not provided in review materials.

Administration

Can be self-administered. Scoring provided by publisher.

Statistical Validation

Not provided with review materials.

Uses

The most appropriate use is for professional development.

Cost

Cost varies from $145 to $185 depending on quantity, processing, and the method selected to capture data.

Leadership Sphere Inventory (LSI)

Development Dynamics
P.O. Box 26026
Austin, TX 78755
Phone: 512-450-1854
Fax: 512-450-1854
Email: Awmsmith@aol.com

Purpose

The LSI is designed to help leaders understand how they view their roles and priorities as leaders.

Target Group

Anyone in a leadership role—executives, administrators, commanders, managers, and supervisors.

Description

The LSI is a self-assessment instrument that seeks to measure how leaders balance the different elements of leadership style. It consists of 24 items that ask participants to rank order a set of berhaviors and beliefs from 1 ("least indicative") to 5 ("most indicative"). In each case, participants are asked to rank the items based on present behavior and potential (ideal) behavior. The questions yield data on the relative strengths of four major styles: interveners, implementers, improvers, and innovators. There is also an "integrated" style with various combinations of the first four. Discrepancies between present behavior and ideal behavior indicate developmental needs.

Feedback

The LSI provides directions for participants to self-score their responses. Numerical and graphic feedback is provided, though the review materials included little narrative explanation.

Followup

The test is structured to provide direction for developmental activity, but the review materials did not include extensive development activities. A separate manual provides more detailed interpretation and applications.

Theory and Rationale

The LSI is based on research and field studies focused on the way that individuals view their relative roles and priorities as leaders.

Administration

The LSI is self-administered and self-scored.

Statistical Validation

Not provided with review materials.

Uses

The LSI is designed for self-evaluation and professional development; it "should not be used to assign or change one's work position or status or to evaluate one's work performance."

Cost

Several options of instruments and packages are available to individualize the LSI for each organization. Inquire at Development Dynamics for details.

Leadership Workstyles

Acumen
4000 Civic Center Drive
5th Floor
San Rafael, CA 94903
Phone: 800-544-8626
Fax: 415-479-5358
Email: acumen@acumen.com
www.acumen.com
Contact: Jeannie Elrod

Purpose

Leadership Workstyles seeks to stimulate leadership development by informing leaders of their characteristic styles.

Target Group

Leaders of organizations in a wide range of industries, including education.

Description

Leadership Workstyles is a multirater instrument that provides feedback on twelve stylistic dimensions: humanistic-helpful, affiliation, approval, conventional, dependence, apprehension, oppositional, power, competition, perfectionism, achievement, and self-actualization. (A version of the instrument may be taken online at Acumen's website.)

Feedback

Participants receive graphic and narrative feedback, with an emphasis on showing how their style helps and hinders four major management functions: managing tasks, managing people, managing conflicts, and influencing/leading others. The report also summarizes perceptions of other raters and provides a comparative analysis.

Followup

The feedback report includes brief suggestions about using the results to increase managerial effectiveness.

Theory and Rationale

Not provided in review materials.

Administration

Can be self-administered. Scoring provided by publisher.

Statistical Validation

Not provided with review materials.

Uses

The most appropriate use is for professional development.

Cost

Cost varies from $145 to $185 depending on quantity, processing, and the method selected to capture data.

Life Styles Inventory (LSI)

Human Synergistics
216 Campus Drive
Arlington Heights, IL 60004
Phone: 847-590-0995
Fax: 847-590-0997
Email: info@hscar.com
www.hscar.com

Purpose

The LSI is designed to increase productivity and develop leadership by helping participants understand their style of thinking, behaving, and interacting.

Target Group

Leaders in all types of organizations.

Description

The Life Styles Inventory is a 360-degree feedback instrument that assesses twelve basic thinking patterns, or styles: humanistic-encouraging, affiliative, approval, conventional, dependent, avoidance, oppositional, power, competitive, perfectionistic, achievement, and self-actualizing. These are further grouped into categories of "constructive," "passive/defensive," or "aggressive/defensive." The inventory contains 240 items in the form of words or short phrases; participants are asked to respond with a "2" if the word is "like you most of the time," "1" if the word is "like you quite often," or "0" if the word is "essentially unlike you."

The "LSI 1" is the self-assessment form; the "LSI 2" is aimed at eliciting feedback from others.

Feedback

Participants receive self-assessment information (LSI 1) in the form of a graphic "circumplex" that charts the relative strengths of each of the twelve styles. In each case, scores are characterized as "high," "medium," or "low," based on a norm group of 9,207 individuals. The LSI 2 adds a similar plotting for responses of others, and also adds fourteen "summary perceptions" showing how others view the person being assessed. Detailed interpretations are provided in self-development guides for both LSI 1 and LSI 2.

Followup

The publisher provides self-development guides for both LSI 1 and LSI 2 that contain extensive development activities, including "thought starters" and "change suggestions" for each style.

Theory and Rationale

The LSI is based on the assumption that behavior is caused by thoughts and self-concept, and therefore it can only be understood by knowing those thoughts.

Administration

Self-administered and self-scored. The LSI 1 takes 20-30 minutes to complete and 10-15 minutes to score; the LSI 2 requires 20-30 minutes to administer and 45-60 minutes to score.

Statistical Validation

Not provided with review materials. A bibliography cites a number of published research studies probing the reliability and validity of the LSI.

Uses

The LSI is most appropriate for professional development.

Cost

The LSI 1 kit, which includes self-inventory, LSI 1 self-development guide, and profile summary cards, is $29. The LSI 2 kit, which includes five "description by others" inventories, LSI 2 self-development guide, scorer's worksheet and instructions, and profile supplement, costs $51. The combined kit for LSI 1 and LSI 2 is $70.

Myers-Briggs Type Indicator (MBTI)

Consulting Psychologists Press
3803 E. Bayshore Road

Palo Alto, CA 94303
Phone: 800-624-1765
Fax: 650-969-8608
www.cpp~db.com

Purpose

The MBTI is designed for use in a wide variety of settings, including business, counseling, and education, for personal and organizational development, team building, counseling, identification of learning styles, and many other activities.

Target Group

Anyone from age 14 to adult.

Description

The MBTI is not specifically a leadership assessment, but it assesses personal traits that affect the ways leaders interact with others and carry out their jobs. Form G (self-scorable) contains ninety-four items that determine preferences in four areas: extraversion-introversion, sensing-intuition, thinking-feeling, and judging-perceiving. These four areas are further combined into sixteen personality "types." Inventory items are forced-choice, asking participants to choose between preferred behaviors or appealing words.

Feedback

Form G (self-scorable) provides brief explanations of the four major areas as well as the sixteen personality types. The publishers provide a variety of more detailed narrative reports as well as advanced versions of the instrument that provide expanded interpretations.

Followup

The publishers offer a number of books and video materials that explore the implications and applications of the MBTI.

Theory and Rationale

The MBTI is based on the personality theory of Carl Jung, who believed that human beings could be categorized into several psychological types, each of which was characterized by certain patterns of thinking and behavior. Through repeated empirical assessments, the MBTI has refined these types into the current configuration.

Administration

Form G (self-scorable) can be self-administered and scored. However, users must meet certain qualifications to license the mate-

rials, including a degree from an acredited college or university and successful completion of a course in the interpretation of psychological tests and measurement at an accredited college or university.

Statistical Validation

Information on validity and reliability is available in the publisher's manual on the MBTI (not reviewed).

Uses

As noted above, the MBTI has many uses. School leaders may find it most useful in understanding how their actions and relationships are affected by their preferred style.

Cost

Form G (self-scorable) for the MBTI is available for $37.50 for a package of ten.

NASSP Assessment Centers

National Association of Secondary School Principals
Office of Professional Development and Assessment
1904 Association Drive
Reston, VA 20191
Phone: 703-860-0200
Fax: 703-476-5432
www.nassp.org

Purpose

The NASSP assessment centers are designed to help school districts identify and develop leadership talent.

Target Group

School leaders or potential school leaders.

Description

The NASSP assessment process uses a variety of simulated leadership tasks to assess the skills of those serving as school leaders or aspiring to school leadership positions. Tasks include group discussions, role plays, inbasket problems, oral presentations, and written papers, with evaluation being done by specially trained observers. Key skills include educational leadership (setting instructional direction, teamwork, and sensitivity); resolving complex problems (judgment, results orientation, and organizational ability); communication

skills (oral communication and written communication); and developing self and others (development of others and understanding own strengths and weaknesses).

The assessment is usually administered in a day-long session at one of over forty assessment centers around the country, followed in several weeks by a feedback session.

In addition to this assessment process ("Selecting and Developing the 21st Century Principal"), NASSP offers two related assessments: Leadership Early Assessment Program (LEAP), aimed at establishing or refining a career-advancement plan for potential or current leaders, and the Superintendent Leadership Development Program (SLDP), designed to help current and potential superintendents construct a development plan (*not* intended for selection purposes).

Feedback

Participants receive written and oral feedback including development options several weeks following the assessment.

Followup

The feedback session includes discussion of development plans. NASSP offers fourteen or fifteen development programs appropriate for following up on the results of the assessment.

Theory and Rationale

The main assessment process ("Selecting and Developing the 21st Century Principal") is newly revised, based on research by the Interstate School Leaders Licensure Consortium (ISLLC), National Policy Board for Educational Administration (NPBEA), and the National Association of Secondary School Principals.

Administration

Assessments are conducted and followed up by trained assessors, typically at a regional assessment center (however, NASSP will provide training for local districts to conduct assessments).

Statistical Validation

Several validation studies have been conducted on the NASSP Assessment Center approach, and others are planned for the future.

Uses

The primary NASSP assessment process ("Selecting and Developing the 21st Century Principal") can be used for selection, promotional readiness, or professional development.

Cost

The cost varies among the regional assessment centers, each of which sets its own fee schedule.

The Principal Perceiver

The Gallup Organization
P.O. Box 5700
Lincoln, NE 68505
Phone: 402-489-9000
Fax: 402-486-6317
www.gallup.com

Purpose

The Principal Perceiver is designed for the selection and development of school leaders.

Target Group

Inservice principals or prospective principals.

Description:

The Principal Perceiver consists of a structured interview conducted and scored by a trained facilitator/assessor. The instrument is designed to identify twelve key "themes" in the beliefs and behaviors of school leaders: commitment, ego drive, achiever, developer, individualized perception, relator, stimulator, team, arranger, command, discipline, and concept.

Feedback

Participants or users receive verbal and written feedback showing scores on each theme and overall score. Graphic feedback shows low and high areas. The scoring is based on "ideal answers" rather than empirical norm.

Followup

None indicated in review materials, though the publishers suggest that the results can be used to identify areas of strength and weakness.

Theory and Rationale

Perceiver interviews are based on the belief that people show spontaneous, recurring patterns of thought, feeling, and behavior that predict how they are likely to perform in a given situation. Through

empirical research, the publishers have established correlations between key indicators and job performance.

Administration

The Principal Perceiver must be administered, scored, and interpreted by a trained and certified assessor. Districts may arrange to have staff members become certified, or the Gallup Organization will provide someone to conduct interviews and feedback by telephone.

Statistical Validation

Not provided with review materials.

Uses

Selection and/or development of K-12 school administrators.

Cost

The Principal Perceiver, which is administered, scored, and interpreted in a structured interview by a trained assessor in the course of a four- to five-day training seminar, costs $2,150 per person.

Professional Development Inventory

National Association of Elementary School Principals
1615 Duke Street
Alexandria, VA 22314
Phone: 703-684-3345
Fax: 703-518-6281
Contact: Merrie Hahn
www.naesp.org

Purpose

The PDI is an individual performance assessment designed to serve as a tool for sharpening understanding of the technical and practical aspects of the principalship. It can be used for diagnosing inservice needs, for preparing for licensing or recertification requirements, and for developing a personal professional-growth plan.

Target Group

The PDI is suitable for elementary or secondary principals.

Description

The PDI is a one-day performance assessment offered at licensed assessment centers affiliated with NAESP. It involves seven exer-

cises designed to "simulate a day in the life of a school principal." The activities require role playing or written responses.

The assessment measures thirteen dimensions of leadership that are grouped under the following five typical responsibilities of principals: instruction, organizational development, supervision and evaluation of staff, communications, and human relations. The thirteen dimensions of school leadership are as follows: planning, organizing, problem-solving, creativity, decisiveness, systems analysis, vision, communications, instructional leadership, group leadership and team building, climate development, moral responsibility, and instructional analysis.

Feedback

Participants generate video and written responses that are evaluated by specially trained and experienced practitioners. Their feedback report identifies specific behaviors that contributed to the scores on each of the thirteen dimensions. Skill levels on each dimension are rated as exceptional, adequate, or inadequate, and the report includes percentile scores measured against a national database of experienced principals.

Exceptional scores are at the eighty-fourth percentile or above, whereas inadequate scores are those falling at the sixteenth percentile or lower. Percentile ratings for specific behavioral descriptors are also included.

The report is available about three weeks after the assessment and is delivered in a one-on-one conference. The participant is the sole recipient of the report.

Followup

The feedback report identifies areas that could benefit from development, with some suggestions for improvement; however, these are not specifically keyed to the PDI. It also lists resources such as seminars, training institutes, bibliographies, AV/computer resources, and projects/activities that can be executed in the organizational context of the school.

Theory and Rationale

The PDI was developed in the 1980s by a team of practitioners and academics at the University of Washington and has been extensively refined since then. The content and methods have been drawn from the effective-schools literature, surveys of practicing school administrators, studies documenting the relationship between self-

awareness and performance enhancement, and school district findings on the effectiveness of training experiences tailored to an initial diagnosis of individual needs.

Administration

The PDI must be taken at designated assessment centers and requires a full day.

Statistical Validation

Content and construct validity have been established for the PDI; reliability is controlled by the licensing requirements and by the process established for administration and scoring. A concurrent validation project is currently under way, and studies of predictive validity will also be carried out.

Uses

The PDI is used primarily as a professional development tool, and is not appropriate for selection or other employment decisions.

Cost

Varies. Districts wishing to have NAESP conduct an assessment onsite must pay for travel, overnight accommodations, and per-diem for two role players and one lead trainer, as well as a $400 honorarium for each role player. In addition, there is a $400 per-participant fee for each simulation conducted and scored by NAESP. Licensed organizations (state principals' associations and universities) may conduct and score their own simulations for a per-participant fee of $170, and may set rates for conducting simulations.

Profile of Aptitude for Leadership (PAL)

Training House, Inc.
P.O. Box 3090
Princeton, NJ 08543
Phone: 609-452-1505
Fax: 609-243-9368

Purpose

PAL is designed to identify and stimulate thinking about leadership style.

Target Group

PAL can be used with leaders in all types of organizations.

Description

The instrument consists of twelve sets of four statements each describing the leader's beliefs or actions. Participants have six points to assign to the four statements in whatever combination best expresses their leadership stance. For example, six points could be given to one statement and no points to the other, or two statements could be given two points each while the others received a single point, and so forth. Results indicate the degree to which a leader operates as manager/administrator, supervisor/coach, entrepreneur/visionary, and technician/specialist.

Feedback

The tests are self-scored and self-recorded. The results are based on the distribution of seventy-two possible points; the higher the score, the stronger the style.

Followup

The feedback form includes basic interpretive comments, along with several reflective questions on each of the four styles.

Theory and Rationale

The basis for the four styles is not specifically addressed in the review materials, though each category is frequently found in theory and research about leadership. PAL assumes that every leader uses all four styles to varying degrees and that all four are necessary.

Administration

PAL can be self-administered and self-scored in less than thirty minutes.

Statistical Validation

Not included with review materials.

Uses

PAL is most appropriately used for development.

Cost

The minimum cost is $80 for a set of 20 packets; additional copies cost $4.00 each for 21 to 99 copies. One hundred copies cost $320, with $3.20 for each additional copy.

The Survey of Leadership Practices

The Clark Wilson Group
1320 Fenwick Lane
Suite 708
Silver Spring, MD 20910
Phone: 800-537-7249
Fax: 301-495-5842
www.cwginc.com

Purpose

The SLP is a multirater assessment designed to give organizational leaders feedback on their efforts to move the organization toward positive change.

Target Group

Organizational leaders, including middle and senior managers, project leaders, and professional people who must build support for their innovations.

Description

The SLP is based on the concept of six "task cycle" skills required for bringing change to organizations: entrepreneurial vision (vision/imagination and risk-taking/venturesomeness); leadership for change (organizational sensitivity and encouraging participation); gaining commitment (teaming/empowering and persuasiveness); monitoring personal impact (feedback); drive (standards of performance, energy, perseverance, and push/pressure); and recognizing performance (sharing credit).

Feedback is also provided on "residual impact" (effectiveness/outcomes, coping with stress, and trustworthiness), as well as "sources of power" (how the leader seeks to influence people).

The inventory contains 85 items on a 7-point Likert scale, ranging from 1 ("never or to a small extent") to 7 ("always or to a great extent"). In addition, three open-ened questions ask what the leader should continue doing, stop doing, and do to become more effective.

Feedback

Participants receive verbal, graphic, and numerical feedback and interpretations for each of the task cycle skills, residual impact, and sources of power. The report shows summary scores for each category, as well as item-by-item breakouts. Self-ratings are compared with those of supervisors, peers, and subordinates. Summary scores

include percentile ratings based on a large sample of managers. Results for open-ended questions are also listed.

Followup

The narrative feedback report includes brief advice for development. A resource guide providing more extensive development assistance is also available (not provided with review materials).

Theory and Rationale

The SLP is based on the idea that leadership depends on skill in accomplishing a sequenced series of tasks, beginning with vision and concluding with recognition of performance. Skill on these tasks will be related to the leader's perceived effectiveness.

Administration

Surveys may be administered in pencil-and-paper form or online. Certification is required to administer the survey and provide feedback.

Statistical Validation

Considerable evidence on validity of the SLP is provided. A research bibliography cites numerous studies using the SLP and closely related instruments.

Uses

Most appropriate for professional development.

Cost

Each survey is $21, which includes the cost of scoring. Quantity discounts are available.

The Survey of Management Practices

The Clark Wilson Group
1320 Fenwick Lane
Suite 708
Silver Spring, MD 20910
Phone: 800-537-7249
Fax: 301-495-5842
www.cwginc.com

Purpose

The SMP is a multirater assessment designed to give organizational leaders feedback on their management skills and practices.

Target Group

Supervisors or others responsible for day-to-day activities of an organization.

Description

The SMP is based on the concept of six "task cycle" skills required for managing organizations effectively: making goals clear and important (clarification of goals and objectives); planning and problem-solving (upward communication, orderly work planning, and expertise); facilitating the work of others (work facilitation); providing feedback; exercising positive control (time emphasis, control of details, goal pressure, delegation); and reinforcing good performance (recognition for good performance).

Feedback is also provided on interpersonal relations (approachability, teambuilding, interest in subordinate growth, and building trust), as well as group motivation and morale (work involvement, coworker competence, team atmosphere, opportunity for growth, tension level, organization climate, general morale, and commitment).

The inventory contains 145 items on a 7-point Likert scale, ranging from 1 ("never or to a small extent") to 7 ("always or to a great extent").

Feedback

Participants receive verbal, graphic, and numerical feedback and interpretations for each of the task cycle skills, interpersonal relations, and group motivation and morale. The report shows summary scores for each category, as well as item-by-item breakouts. Self-ratings are compared with those of supervisors, peers, and subordinates. Summary scores include percentile ratings based on a large sample of managers.

Followup

The narrative feedback report includes brief advice for development. A resource guide providing more extensive development assistance is also available (not provided with review materials).

Theory and Rationale

The SMP is based on the idea that effective management depends on skill in accomplishing a sequenced series of tasks, beginning with goal-setting and concluding with recognition for good performance. The SMP assumes that effective managers are those who balance structure with consideration.

Administration

Surveys may be administered in pencil-and-paper form or online. Certification is required to administer the survey and provide feedback.

Statistical Validation

Considerable evidence on validity of the SMP is provided. A research bibliography cites numerous studies using the SMP and closely related instruments.

Uses

Most appropriate for professional development.

Cost

Each survey is $21, which includes the cost of scoring. Quantity discounts are available.

Bibliography

Many of the items in this bibliography are indexed in ERIC's monthly catalog *Resources in Education* (*RIE*). Reports in *RIE* are indicated by an "ED" number. Journal articles, indexed in ERIC's companion catalog, *Current Index to Journals in Education*, are indicated by an "EJ" number.

Most items with an ED number are available from ERIC Document Reproduction Service (EDRS), 7420 Fullerton Rd., Suite 110, Springfield, VA 22153-2852.

To order from EDRS, specify the ED number, type of reproduction desired—microfiche (MF) or paper copy (PC), and number of copies. Add postage to the cost of all orders and include check or money order payable to EDRS. For credit card orders, call 1-800-443-3742.

Alsop, Joseph, and Stuart Alsop. "Taft and Vandenberg." *Life* 21,15 (October 7, 1946): 102-14.

American Educational Research Association, American Psychological Association, and National Council on Educational Measurement. *Standards for Educational and Psychological Testing*. Washington, D. C.: American Psychological Association, 1985.

Anderson, Mark. *Principals: How To Train, Recruit, Select, Induct, and Evaluate Leaders for America's Schools*. Eugene, Oregon: ERIC Clearinghouse on Educational Management, University of Oregon. 1991. ED 337 843.

Bass, Bernard M. *Bass & Stodill's Handbook of Leadership: Theory, Research, and Managerial Applications*. 3rd edition. New York: The Free Press, 1990.

Bass, Bernard M., and Bruce J. Avolio. "Introduction." In *Improving Organizational Effectiveness Through Transformational Leadership*. Edited by Bernard Bass and Bruce Avolio. 1-9. Thousand Oaks, California: Sage Publications, 1994.

Benfari, Robert. *Understanding Your Management Style: Beyond the Myers-Briggs Type Indicators*, Lexington, Massachusetts: Lexington Books, 1991.

Bennis, Warren. *Managing People Is Like Herding Cats*. Provo, Utah: Executive Excellence Publishing, 1997.

_____. "Transformative Power and Leadership." In *Leadership and Organizational Culture: New Perspectives on Administrative Theory and Practice*, edited by Thomas J. Sergiovanni and John E. Corbally, 64-71. Urbana, Illinois: University of Illinois Press, 1984.

Bolman, Lee, and Terry Deal. *Reframing Organizations: Artistry, Choice and Leadership*. San Francisco: Jossey-Bass, 1991. 492 pages. ED 371 457.

Blumberg, Arthur. *School Administration as a Craft: Foundations of Practice*. Boston: Allyn and Bacon, 1989. ED 304 773.

Chappelow, Craig. In *The Center for Creative Leadership* Handbook of Leadership Development, edited by Cynthia McCauley and others. Greensboro, North Carolina: Center for Creative Leadership, 1998. EA 029 332.

Chemers, Martin M. *An Integrated Theory of Leadership*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 1997.

Clark, Kenneth E., and Miriam Clark. *Choosing To Lead*. Second Edition. Greensboro, North Carolina: Center for Creative Leadership, 1996. ED 396 435.

Coutts, J. Douglas. "Why Principals Fail: Are National Professional Standards Valid Measures of Principal Performance?" *ERS Spectrum* (Fall 1997): 20-24.

Dalton, Maxine. "Best Practices: Five Rationales for Using 360-Degree Feedback in Organizations." In *Maximizing the Value of 360-Degree Feedback*. Edited by Walter W. Tornow and colleagues. 59-77. San Francisco: Jossey-Bass, 1998.

The Education Commission of the States. *Designing and Implementing Standards-Based Accountability Systems*. Denver, Colorado:1998.

Freeman, Frank H; Katherine B. Knott; and Mary K. Schwartz. *Leadership Education: A Source Book, 1996-1997*. 6th ed., volume 2. Greensboro, North Carolina: Center for Creative Leadership, 1996.

Greenfield, William. "Moral, Social, and Technical Dimensions of the Principalship." *Peabody Journal of Education* 63, 1 (Fall 1986): 130-49. EJ 354 953.

Hersey, Paul, and Kenneth Blanchard. *Management of Organizational Behavior: Utilizing Human Resources*. Sixth edition. Englewood Cliffs, New Jersey: Prentice-Hall, 1993.

Hoyle, John R.; Fenwick W. English; and Betty E. Steffy. *Skills for Successful 21st Century School Leaders: Standards for Peak Performers.* Arlington, Virginia: American Association of School Administrators, 1998.

Joint Committee on Testing Practices. *Code of Fair Testing Practices in Education.* Washington, D.C.: 1988. 5 pages. ED 301 574.

Kaplan, Robert E., and Charles J. Paulus. *Enhancing 360-Degree Feedback for Senior Executives: How to Maximize the Benefits and Minimize the Risks.* Greensboro, North Carolina: Center for Creative Leadership, 1994. 31 pages. ED 382 803.

Klenke, Karin. *Women and Leadership: A Contextual Perspective.* New York: Springer Publishing Company, 1996.

Kouzes, James, and Barry Posner. *Credibility: How Leaders Gain and Lose It, Why People Demand It.* San Francisco: Jossey-Bass, 1993.

Leithwood, Kenneth; Paul T. Bedley; and J. Bradley Cousins. *Developing Expert Leadership for Future Schools.* London: Falmer Press, 1994.

Lepsinger, Richard, and Antoinette D. Lucia. *The Art and Science of 360° Feedback.* San Francisco: Jossey-Bass, 1997.

Manatt, Richard P. "Feedback from 360 Degrees: Client-Driven Evaluation of School Personnel." *School Administrator* 54, 3 (March 1997): 8-13. EJ 540 889.

Myers, Isabel Briggs, and Peter B. Myers. *Gifts Differing.* Palo Alto, California: Consulting Psychologists Press, 1980.

Sergiovanni, Thomas J. *Leadership for the Schoolhouse: How Is It Different? Why Is It Important?* San Francisco: Jossey-Bass, 1996. 203 pages. ED 390 153.

Sharp, William L.; James K. Walter; and Helen M. Sharp. *Case Studies for School Leaders: Implementing the ISLLC Standards.* Lancaster, Pennsylvania: Technomic Publishing, 1998.

Van Velsor, Ellen. "Designing 360-Degree Feedback To Enhance Involvement, Self-Determination, and Commitment." In *Maximizing the Value of 360-Degree Feedback*, edited by Walter W. Tornow and colleagues. 149-95. San Francisco: Jossey-Bass, 1998.

CURRENT TITLES

ERIC Clearinghouse on Educational Management

___ *Measuring Leadership: A Guide to Assessment for Development of School Executives*, by Larry Lashway @$9.75

___ *Student Motivation: Cultivating a Love of Learning*, by Linda Lumsden @ $9.50

___ *Leading with Vision*, by Larry Lashway @ $13.50

___ *School Leadership: Handbook for Excellence*, 3rd ed., edited by Stuart C. Smith and Philip K. Piele
___ copies of Cloth @ $29.95
___ copies of Paper@ $19.95

___ *Roadmap to Restructuring: Charting the Course of Change in American Education*, 2nd ed., by David T. Conley
___ copies of Cloth @ $34.95
___ copies of Paper@ $23.95

___ *Learning Experiences in School Renewal: An Exploration of Five Successful Programs*, edited by Bruce Joyce and Emily Calhoun @ $14.50

___ *Transforming School Culture: Stories, Symbols, Values, and the Leader's Role*, by Stephen Stolp and Stuart C. Smith @ $12.50

___ *Planning for Effective Staff Development: Six Research-Based Models*, by Meredith D. "Mark" Gall and Roseanne O'Brien Vojtek @ $6.95

___ *Implementing Problem-Based Learning in Leadership Development*, by Edwin M. Bridges and Philip Hallinger @ $14.95

___ *On Understanding Variables & Hypotheses in Scientific Research*, by W.W. Charters, Jr. @ $6.95

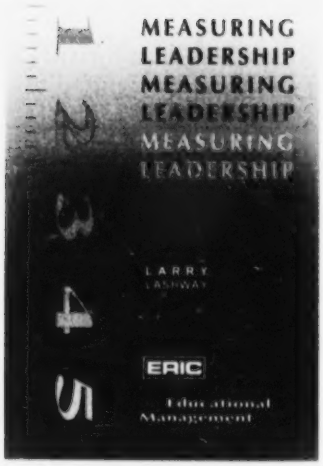
How to Order: You may place an order by sending a check or money order, mailing or faxing a purchase order, or calling with a Visa or MasterCard number. Add 10% for S & H (minimum $4.00). Make payment to **University of Oregon/ERIC** and mail to ERIC/CEM, 5207 University of Oregon, Eugene, Oregon 97403-5207. Shipping is by UPS ground or equivalent.

Telephone (800) 438-8841

Fax (541) 346-2334

You can also order online (with Visa or MasterCard) from our website—your gateway to information about educational policy and management.

http://eric.uoregon.edu



Larry Lashway guides district administrators and principals through the maze of issues on leadership assessment: what leadership is, how to select an instrument, and how to interpret and use the data it generates. Descriptions of about twenty instruments are included.

The book focuses less on detailed descriptions of instruments than on a *process* that begins with reflection on the district's leadership needs and ends when participants begin to act on the implications of the results.

Contents

The Measure of a Leader

What Do Leadership Tests Measure?

Selecting the Best Instrument

Fulfilling the Promise of Assessment

Instrument Profiles

Kenneth Leithwood says:

"Written primarily for those with districtwide responsibilities, this is a 'how to' book for school-leader selection, appraisal, and development. It is written extremely well with this audience in mind. And while the author touches, of necessity, on only a sample of the total set of leadership instruments that are available, he provides quite useful guidance for district administrators no matter which leadership instruments they select for their own purposes." (From the Foreword)

BEST COPY AVAILABLE

ERIC Clearinghouse on Educational Management

ISBN 0-86552-140-9
50975